



## Peripheral auditory processing and speech reception in impaired hearing

Strelcyk, Olaf

*Publication date:*  
2009

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Strelcyk, O. (2009). *Peripheral auditory processing and speech reception in impaired hearing*. Technical University of Denmark. CONTRIBUTIONS TO HEARING RESEARCH No. 2

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

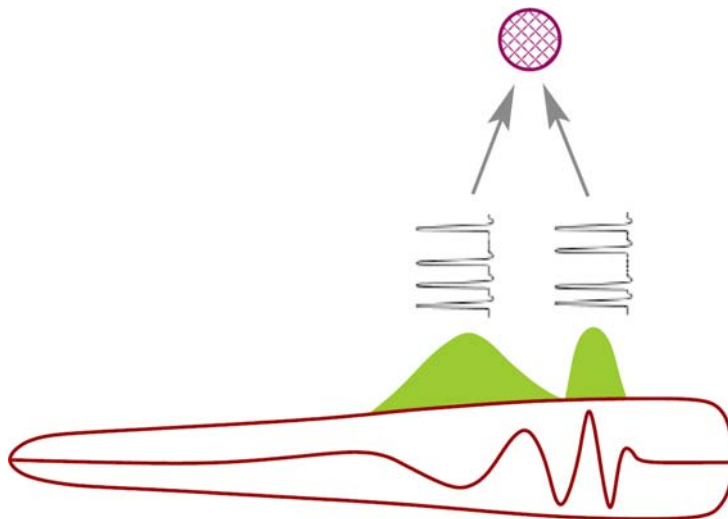
CONTRIBUTIONS TO  
HEARING RESEARCH

Volume 2

---

*Olaf Strelcyk*

# Peripheral auditory processing and speech reception in impaired hearing

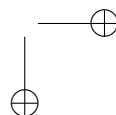
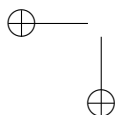
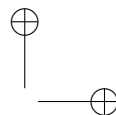
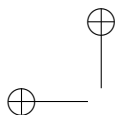


# Peripheral auditory processing and speech reception in impaired hearing

PhD thesis by  
Olaf Strelcyk



Technical University of Denmark  
2009



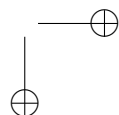
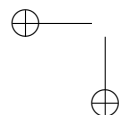
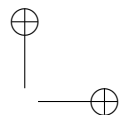
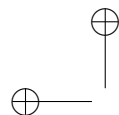
---

## Abstract

---

One of the most common complaints of people with impaired hearing concerns their difficulty with understanding speech. Particularly in the presence of background noise, hearing-impaired people often encounter great difficulties with speech communication. In most cases, the problem persists even if reduced audibility has been compensated for by hearing aids. It has been hypothesized that part of the difficulty arises from changes in the perception of sounds that are well above hearing threshold, such as reduced frequency selectivity and deficits in the processing of temporal fine structure (TFS) at the output of the inner-ear (cochlear) filters. The purpose of this work was to investigate these aspects in detail.

One chapter studies relations between frequency selectivity, TFS processing, and speech reception in listeners with normal and impaired hearing, using behavioral listening experiments. While a correlation was observed between monaural and binaural TFS-processing deficits in the hearing-impaired listeners, no relation was found between TFS processing and frequency selectivity. TFS processing was correlated with speech reception in background noise. Two following chapters investigate cochlear response time (CRT) as an important aspect of the cochlear response to incoming sounds, using objective and behavioral methods. Alterations in CRT were observed for hearing-impaired listeners. A good correspondence between objective and behavioral estimates of CRT indicated that a behavioral lateralization method may be useful for studying spatiotemporal aspects of the cochlear response in human listeners. Behaviorally estimated filter bandwidths accounted for the observed alterations of CRTs in the hearing-impaired listeners, i.e., CRT was found to be inversely related to individual filter bandwidth. Overall, this work provides insights into factors affecting auditory processing in listeners with impaired hearing and may have implications for future models of impaired auditory signal processing as well as advanced compensation strategies.



---

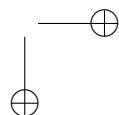
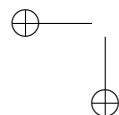
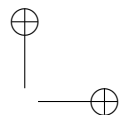
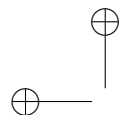
## Resumé

---

En af de mest almindelige klager blandt mennesker med nedsat hørelse er deres vanskeligheder med at forstå tale og problemerne optræder især når der er baggrundsstøj til stede. I de fleste tilfælde fortsætter problemerne selv om høretabet bliver kompenseret med et høreapparat. En hypotese er at en del af disse vanskeligheder skyldes en ændret opfattelse af lyde der ligger langt fra høretærsklen. Dette kunne være reduceret frekvensselektivitet og nedsat evne til at bearbejde den temporale finstruktur (TFS) der optræder på udgangen af filtrene i det indre øre. Formålet med dette projekt var at foretage en detaljeret undersøgelse af disse forhold.

Et kapitel omtaler en undersøgelse af sammenhængen mellem frekvensselektivitet, TFS-behandlingsevne og opfattelse af tale hos personer med normal og med nedsat hørelse ved hjælp af lytteforsøg. Der blev observeret en sammenhæng mellem nedsat monaural og binaural TFS-behandlingsevne hos personer med høretab. Derimod blev der ikke fundet nogen sammenhæng mellem TFS-behandlingsevne og frekvensselektivitet. TFS-behandlingsevnen var korreleret med forståelsen af tale i baggrundsstøj. To kapitler omtaler en undersøgelse af det indre øres (cochlea) responstid (CRT) som reaktion på indkommende lyd. Hos personer med høretab blev der fundet afvigelser i CRT i forhold til normalthørende. Der var god overensstemmelse mellem estimer af CRT baseret på en objektiv og en subjektiv målemetode. Den subjektive metode kunne være nyttig til at belyse rumlige og tidsmæssige aspekter af CRT hos mennesker. Filterbåndbredder estimeret på basis af lytteforsøg korrelerede med de observerede afvigelser i CRT hos personer med høretab idet CRT var omvendt proportional med de individuelle filterbåndbredder.

Samlet set, giver dette arbejde indsigt i de forhold der har indflydelsen på lydopfattelsen hos personer med høretab og kan have betydning for fremtidige auditive modeller til beskrivelse af høretab. Arbejdet kan desuden have betydning for udvikling af nye avancerede strategier til at kompensere for høretab.





---

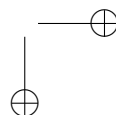
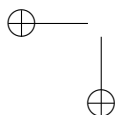
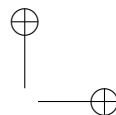
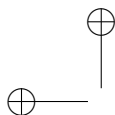
## Zusammenfassung

---

Häufig klagen schwerhörige Menschen über Probleme beim Sprachverstehen. Vor allem Kommunikation im Störgeräusch stellt für sie eine große Herausforderung dar. Selbst wenn der Hörverlust durch Hörgeräte kompensiert wird, d.h. Hörbarkeit wiederhergestellt wird, bleiben die Probleme oft bestehen. Die vorliegende Arbeit untersucht die Hypothese, dass ein Teil der Probleme auf Veränderungen in der Hörwahrnehmung oberhalb der Hörschwelle zurückzuführen ist, insbesondere auf verminderte Frequenzselektivität und Defizite in der Verarbeitung der temporalen Feinstruktur (TFS) am Ausgang der Frequenzgruppen des Innenohres (Kochlea).

Im ersten Kapitel wird mittels perzeptiver Hörexperimente der Zusammenhang zwischen Frequenzselektivität, TFS-Verarbeitungsvermögen und Sprachverstehen bei Normalhörenden sowie Schwerhörigen untersucht. Während ein Zusammenhang zwischen einohrigem und beidohrigem Verarbeitungsvermögen von TFS bei Schwerhörigen gefunden wurde, gab es keine Relation zwischen Frequenzselektivität und TFS-Verarbeitungsvermögen. Es zeigten sich allerdings Korrelationen zwischen TFS-Verarbeitung und Sprachverstehen im Störgeräusch. Zwei weitere Kapitel befassen sich mit der Antwortzeit des Innenohres auf auditorische Stimuli. Hier wurden sowohl perzeptive als auch objektive Messmethoden verwendet. Bei den schwerhörigen Versuchspersonen wurden Abweichungen in der kochlearen Antwortzeit gegenüber normalhörenden Versuchspersonen festgestellt. Darüber hinaus wurde eine gute Übereinstimmung zwischen perzeptiven und objektiven Schätzwerten der kochlearen Antwortzeit gefunden. Somit könnte die perzeptive Methode dazu genutzt werden, räumliche und zeitliche Aspekte der kochlearen Verarbeitung beim Menschen zu untersuchen. Die beobachteten Abweichungen in der kochlearen Antwortzeit bei Schwerhörigen konnten auf Unterschiede in der Bandbreite der Frequenzgruppen zurückgeführt werden. Die individuelle kochleare Antwortzeit verhielt sich umgekehrt proportional zur perzeptiv ermittelten Frequenzgruppenbandbreite.

Zusammenfassend trägt diese Arbeit zu einem tieferen Verständnis von Faktoren der auditorischen Signalverarbeitung bei, welche die Hörwahrnehmung von schwerhörigen Menschen beeinträchtigen können. Sie hat Konsequenzen sowohl für zukünftige Modelle des auditorischen Systems als auch für moderne Strategien zur Kompensation von Schwerhörigkeit.



---

## Preface

---

Arriving at the institute in the morning—just on the steps—the smell of the sea. What a blessing! Three and a half years of living and working in Copenhagen have been a rich experience, and I would not want to do without it. I would like to express my gratefulness to...

Torsten Dau, who has supervised this work. Thank you, Torsten, for introducing me to the field, trusting in me, reassuring me, and giving me plenty of freedom. Your straightforward way made it an ease and a great pleasure to work with you!

Graham Naylor, who has co-supervised part of this work, Christian Lorenzi, who welcomed me for a month's stay in Paris, Brian C. J. Moore, Steven Greenberg, Eric R. Thompson, and Sabine Caminade for valuable comments and stimulating discussions.

All of my friends and colleagues at the centre and the institute, for your helping hands, your commitment—personal and professional—and every smile in the hallway! In particular, I would like to thank Brent C. Kirkwood, James Harte, Claus F. Christiansen, Dimitrios Christoforidis, and Torben Poulsen for valuable comments on previous versions of this manuscript.

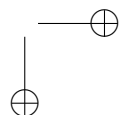
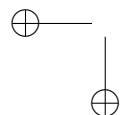
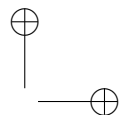
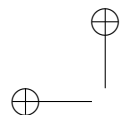
All the listeners for participation in hours of testing.

The Danish Research Foundation, the Danish Graduate school SNAK “Sense organs, neural networks, behavior, and communication”, and the Oticon Foundation, for the funding of this work.

Friends. Dear friends—you have given me such treasure!

My parents and family, for loving and backing me all the way.

*Olaf Strelcyk, April 3, 2009*



---

# Contents

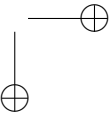
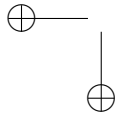
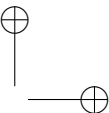
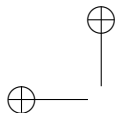
---

<b>List of abbreviations</b>	<b>xv</b>
<b>1 General introduction</b>	<b>1</b>
<b>2 Relations between frequency selectivity, temporal fine-structure processing, and speech reception</b>	<b>5</b>
2.1 Introduction . . . . .	6
2.2 Methods . . . . .	10
2.2.1 Listeners . . . . .	10
2.2.2 Apparatus . . . . .	12
2.2.3 Statistical analyses . . . . .	13
2.3 Speech reception . . . . .	13
2.3.1 Method . . . . .	13
2.3.2 Results and discussion . . . . .	14
2.4 Frequency selectivity . . . . .	18
2.4.1 Method . . . . .	18
2.4.2 Results and discussion . . . . .	19
2.5 Binaural masked detection . . . . .	21
2.5.1 Method . . . . .	21
2.5.2 Results and discussion . . . . .	21
2.6 Lateralization . . . . .	23
2.6.1 Method . . . . .	23
2.6.2 Results and discussion . . . . .	24

2.7	Frequency modulation detection . . . . .	28
2.7.1	Method . . . . .	28
2.7.2	Results and discussion . . . . .	30
2.8	Comparison of results across tests . . . . .	33
2.8.1	Hearing-impaired listeners . . . . .	33
2.8.2	Listeners with obscure dysfunction . . . . .	41
2.9	Possible underlying impairment mechanisms . . . . .	42
2.10	Summary . . . . .	44
<b>3</b>	<b>Estimation of cochlear response times using lateralization of frequency-mismatched tones</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Auditory brainstem responses . . . . .	48
3.2.1	Method . . . . .	48
3.2.2	Results . . . . .	51
3.3	Lateralization of mismatched tones . . . . .	52
3.3.1	Method . . . . .	52
3.3.2	Results and discussion . . . . .	55
3.4	Overall discussion . . . . .	61
<b>4</b>	<b>Relation between derived-band auditory brainstem response latencies and frequency selectivity</b>	<b>65</b>
4.1	Introduction . . . . .	65
4.2	Auditory brainstem responses . . . . .	67
4.2.1	Method . . . . .	67
4.2.2	Results and discussion . . . . .	71
4.3	Frequency selectivity . . . . .	78
4.3.1	Method . . . . .	78
4.3.2	Results and discussion . . . . .	80
4.4	Relation between cochlear response time and frequency selectivity . .	81

---

4.5	Predicting frequency selectivity from derived-band ABR latencies . .	85
4.5.1	Background . . . . .	85
4.5.2	Method . . . . .	85
4.5.3	Results and discussion . . . . .	86
4.6	Summary . . . . .	88
<b>5</b>	<b>General discussion</b>	<b>89</b>
	<b>Bibliography</b>	<b>93</b>





---

## List of abbreviations

---

<i>nI-mAFC</i>	<i>n</i> -interval, <i>m</i> -alternative, forced choice
ABR	Auditory brainstem response
AM	Amplitude modulation
AN	Auditory nerve
ANOVA	Analysis of variance
B&K	Brüel & Kjær
BM	Basilar membrane
CF	Characteristic frequency
CI	Confidence interval
CRT	Cochlear response time
ERB	Equivalent rectangular bandwidth
FIR	Finite impulse response
FM	Frequency modulation
FMDT	Frequency-modulation detection threshold
HI	Hearing impaired
HL	Hearing level
IHC	Inner hair cell
IPD	Interaural phase difference
ITD	Interaural time difference
LATSSN	Lateralized speech-shaped noise

MEM	Mixed-effects model
MLD	Masking level difference
MULTITALK	Multi-talker babble
NH	Normal hearing
NM	Not measurable
OAE	Otoacoustic emission
OD	Obscure dysfunction
OHC	Outer hair cell
ppe SPL	Peak-to-peak equivalent sound pressure level
PTA	Pure-tone average threshold at 0.5, 1, 2, and 4 kHz
PTA <sub>w</sub>	Locally weighted pure-tone average threshold
RAM	Randomly amplitude modulated
rms	Root-mean-square
roex	Rounded exponential
SAM	Sinusoidally amplitude modulated
SD	Standard deviation
SL	Sensation level
SNR	Signal-to-noise ratio
SPL	Sound pressure level
SRT	Speech reception threshold
SSN	Speech-shaped noise
TFS	Temporal fine structure
TW	Traveling wave
TWOTALK	Two-talker background

# 1

---

## General introduction

---

In 2001, the World Health Organization estimated that, worldwide, 250 million people, i.e., approximately 4% of the global population, have disabling hearing difficulties. In an extensive epidemiological study in Great Britain, Davis (1989) estimated that about 16% of the adult population (17–80 years) have a bilateral hearing loss of at least 25 dB (averaged over the speech frequencies 0.5, 1, 2, and 4 kHz). The (age-adjusted) prevalence of hearing impairment has been found to be on the increase, presumably due to increasing noise exposure and ototoxic drug use (Wallhagen et al., 1997; World Health Organization, 1997). Also, hearing impairment is considered to be of growing importance in view of the aging of populations. Most conductive defects in the transmission of sound to the inner ear can nowadays be successfully rehabilitated by means of surgery. Therefore, the majority of inoperable hearing impairments are sensorineural, i.e., result from defects in the inner ear, auditory nerve, or higher centers of the brain.

Hearing impairment is a communicative handicap: hearing-impaired (HI) people often experience great difficulty with speech communication. These difficulties are typically most pronounced when background noise is present, for example in reverberant environments or in situations with multiple interfering sound sources. Normal-hearing (NH) listeners are able to listen to, and follow, one speaker in the midst of background chatter, a situation referred to as the “cocktail party problem” (Cherry, 1953). Despite the seemingly effortless and intuitive nature of this ability, little is known about the underlying auditory signal processing. It is unclear how the (normally functioning) auditory system parses acoustic scenes to form mental representations of the sound sources. Even less is understood about the factors and mechanisms that are responsible for the reduced performance of HI listeners. While audibility has been shown to be the main determinant of speech reception in quiet, it does not ac-

count to the same degree for speech reception in noise (e.g., Plomp, 1978; Dreschler and Plomp, 1985; Glasberg and Moore, 1989). Other impairment factors besides reduced audibility must be involved. Consequently, for many HI listeners, the problem persists even if reduced audibility has been compensated for by hearing aids.

The ability to hear one sound in the presence of other sounds depends crucially on the auditory system's frequency resolution, or frequency selectivity, which is generally attributed to the inner-ear (cochlear) filters. Typically, HI listeners show reduced frequency selectivity, i.e., they are more susceptible to masking from remote frequency components. Therefore, previous studies have investigated the role of frequency selectivity and found that reduced frequency selectivity could partly account for the HI listeners' problems with speech reception in noise (e.g., Festen and Plomp, 1983; Dreschler and Plomp, 1985; Horst, 1987; van Schijndel et al., 2001).

The signal at the output of the cochlear filters can be considered as a time-varying envelope superimposed on the more rapid fluctuations of a carrier, sometimes called the temporal fine structure (TFS). It has commonly been assumed that mainly envelope cues govern speech reception (for a review, see Lorenzi and Moore, 2008). Recently, however, the processing of TFS information has received considerable attention with regard to speech reception in HI listeners (e.g., Buss et al., 2004; Lorenzi et al., 2006; Hopkins et al., 2008). It has been suggested that deficits in TFS coding might account for the limited ability of HI listeners to take advantage of amplitude fluctuations in a noise background, i.e., to listen in the dips of a fluctuating interferer (e.g., Qin and Oxenham, 2003; Lorenzi et al., 2006). Furthermore, Zeng et al. (2005) suggested that TFS information might be utilized in talker identification and separation. Hence, deficits in TFS coding might account for part of the difficulties with speech reception experienced by HI listeners in complex listening situations.

The cochlea responds to sound with a “traveling wave” that propagates on the basilar membrane within the cochlea (e.g., Ruggero, 1994; Robles and Ruggero, 2001). Several studies have suggested that the extraction of spatiotemporal information, i.e., the combination of phase-locked responses and systematic frequency-dependent delays along the cochlea (associated with the traveling wave), may be important for the decoding of TFS information (e.g., Schroeder, 1977; Loeb et al., 1983; Deng and Geisler, 1987; Shamma et al., 1989). It has been proposed that a distorted

spatiotemporal cochlear response might be, at least partly, responsible for the HI listeners’ deficits in the processing of TFS information (e.g., Moore, 1996; Moore and Skrodzka, 2002). Hence, it would be important to gain a better understanding of how hearing impairment affects the spatiotemporal response pattern. However, so far, empirical evidence for spatiotemporal information processing in humans is lacking, since cochlear response patterns are difficult to monitor.

The purpose of the behavioral and objective experiments presented in this thesis was to explore, in depth, deficits in frequency selectivity and TFS processing in sensorineurally HI listeners and to study the contributions of TFS cues to speech reception in NH and HI listeners. Another goal was to investigate possible alterations of the spatiotemporal cochlear response in HI listeners when compared with NH listeners.

**Chapter 2** of this thesis examines relations between frequency selectivity, monaural and binaural TFS processing, and speech reception in NH and HI listeners, using behavioral methods. Also, listeners with an obscure dysfunction were included, who showed normal audiograms but complained about difficulties understanding speech in noisy backgrounds. Frequency selectivity was assessed, since the contribution of reduced frequency selectivity to observed TFS deficits remained unclear in previous studies (Lorenzi et al., 2006; Hopkins et al., 2008). The results were expected to provide insights into the nature of TFS deficits in HI listeners and into consequences of TFS deficits for speech reception. This could have implications for compensation strategies such as hearing aids or cochlear implants.

The two following chapters investigate cochlear response times (CRTs), which represent an important component of spatiotemporal-processing concepts. **Chapter 3** evaluates the validity of a behavioral method for estimation of CRT disparities between remote places on the basilar membrane and traveling-wave velocities, proposed by Zerlin (1969). The paradigm relies on lateralization of pulsed tones that are interaurally mismatched in frequency. The behavioral estimates were compared with objective estimates for the same NH listeners, based on derived-band auditory brain-stem response (ABR) latencies.

**Chapter 4** examines derived-band ABRs for NH and HI listeners, in order to explore possible alterations in CRT due to hearing impairment. For the same listeners,

behavioral estimates of frequency selectivity were obtained, to study the relationship between CRT and frequency tuning. It was expected that larger across-listener variability within the group of HI listeners, compared to that for the NH listeners, would provide valuable information when investigating this relation. The results might contribute to a better understanding of how hearing impairment affects the cochlear response pattern in human listeners. Furthermore, they might demonstrate a possibility for predicting individual frequency selectivity from objective ABR measurements.

Finally, **Chapter 5** summarizes the main outcomes of this work and discusses its possible implications, for auditory modeling as well as for advanced compensation strategies.

## 2

---

### **Relations between frequency selectivity, temporal fine-structure processing, and speech reception\***

---

Frequency selectivity, TFS processing, and speech reception were assessed for six NH listeners, ten sensorineurally HI listeners with similar high-frequency losses, and two listeners with an obscure dysfunction. TFS processing was investigated at low frequencies in regions of normal hearing, through measurements of binaural masked detection, tone lateralization, and monaural frequency modulation (FM) detection. Lateralization and FM detection thresholds were measured in quiet and in background noise. Speech reception thresholds were obtained for full-spectrum and lowpass-filtered sentences with different interferers. Both the HI listeners and the listeners with obscure dysfunction showed poorer performance than the NH listeners in terms of frequency selectivity, TFS processing, and speech reception. While a correlation was observed between the monaural and binaural TFS-processing deficits in the HI listeners, no relation was found between TFS processing and frequency selectivity. The effect of noise on TFS processing was not larger for the HI listeners than for the NH listeners. Finally, TFS-processing performance was correlated to speech reception in a two-talker background and lateralized noise, but not in amplitude-modulated noise. The results provide constraints for future models of impaired auditory signal processing.

---

\* This chapter is based on Strelcyk and Dau (2009a).

## 6 2. Relations between frequency selectivity, TFS processing, and speech reception

### 2.1 Introduction

Hearing-impaired people often experience great difficulty with speech communication when background noise is present. While audibility has been shown to be the main determinant of speech reception in quiet, it does not account to the same degree for speech reception in noise (e.g., Plomp, 1978; Dreschler and Plomp, 1985; Glasberg and Moore, 1989; Peters et al., 1998). Other impairment factors besides reduced audibility must be involved.

Relations between frequency selectivity and speech reception, particularly in noise, have been reported previously (e.g., Festen and Plomp, 1983; Dreschler and Plomp, 1985; Horst, 1987; van Schijndel et al., 2001). Recently, also the processing of TFS information has received considerable attention with regard to speech reception (e.g., Tyler et al., 1983; Buss et al., 2004; Lorenzi et al., 2006; Hopkins et al., 2008; Lorenzi et al., 2009). While envelope cues are sufficient to achieve good speech reception in quiet (e.g., Shannon et al., 1995), TFS cues may be required to ensure good speech reception in noise (e.g., Nie et al., 2005; Lorenzi and Moore, 2008). In particular, it has been suggested that deficits in TFS coding might account for the limited ability of HI listeners to take advantage of amplitude fluctuations in a noise background, i.e., to listen in the dips of a fluctuating interferer (e.g., Qin and Oxenham, 2003; Lorenzi et al., 2006; Gnansia et al., 2008; Hopkins and Moore, 2009). However, the large variability of performance that is commonly observed across HI listeners makes it difficult to compare results across studies. Hence, only limited conclusions can be drawn about the relations between the different auditory functions, such as frequency selectivity and the processing of TFS. Also the relation between the deficits observed in monaural and binaural TFS processing remains unclear. Knowledge of these relations might shed light on the actual mechanisms and sites of the impairments.

Therefore, in the present study, individual performance on frequency selectivity, monaural and binaural TFS processing, and speech reception was measured using a common set of listeners. This is a similar concept to that used in the studies of Hall et al. (1984) and Gabriel et al. (1992), who examined binaural performance in individual HI listeners. Since the primary objective of the present study was to investigate impairment factors beyond audibility, ten HI listeners with similar high-frequency



## 2.1 Introduction

7

hearing losses were selected to provide a homogeneous group in terms of audibility. In this way, confounding effects of audibility were minimized and more direct conclusions could be drawn from a relatively small number of subjects about possible relations between the tested auditory functions. On the flip side, however, this group of HI listeners represents one homogeneous subset of the overall HI population and therefore one should act with caution in generalizing the results.

Besides the HI listeners, two further subjects were included in the present study. Despite normal audiograms, these subjects complained about difficulties with speech reception in noisy backgrounds. In literature, different terms have been used to refer to this phenomenon: auditory disability with normal hearing (King and Stephens, 1992), obscure auditory dysfunction (Saunders and Haggard, 1989), and King-Kopetzky syndrome (Hinchcliffe, 1992). For simplicity, in the following, these subjects are referred to as having an obscure dysfunction (OD). In view of the heterogeneity of the clinical group of OD patients (e.g., Saunders and Haggard, 1989; Zhao and Stephens, 2000), these two listeners cannot constitute a representative sample and therefore should be regarded as cases. The comparison of performance between the two OD listeners and the HI listeners may provide valuable information on the nature of the underlying impairments in both groups.

Speech reception thresholds (SRTs) for full-spectrum and lowpass-filtered speech were measured in different diotic and dichotic interferers. The other psychoacoustic tests in this study were designed to examine basic auditory functions, mainly at a frequency of 750 Hz. Low-frequency information has been shown to play a dominant role both for monaural abilities, such as the perception of pitch of complex tones (e.g., Terhardt, 1974; Moore et al., 1985), and for binaural abilities such as sound localization (e.g., Wightman and Kistler, 1992). Therefore, the frequency of 750 Hz was chosen to investigate the potential impact of a hearing impairment on auditory processing at low frequencies, even if a hearing loss in terms of elevated audiometric thresholds was present only at higher frequencies. As a basic auditory function, frequency selectivity was estimated via the notched-noise paradigm in simultaneous masking (e.g., Patterson and Nimmo-Smith, 1980).

Throughout the present study, the terms TFS information and TFS processing refer to the temporal fine structure at the output of the cochlear filters. This fine

## 8 2. Relations between frequency selectivity, TFS processing, and speech reception

structure evokes phase-locked activity, i.e., synchronized timing of action potentials, in the subsequent stages of neural processing (see Ruggero, 1992). Apart from phase locking, TFS information may also be coded in terms of a conversion from frequency modulation to amplitude modulation (FM-to-AM) on the cochlear filter skirts, as has been suggested for the detection of high-rate FM (Zwicker, 1956; Moore, 2003). In the present study, however, the focus lies on TFS processing based on phase locking, rather than on the FM-to-AM conversion mechanism.

Evidence for TFS-processing deficits in HI listeners has been found in previous studies of monaural as well as binaural auditory functions. In terms of binaural processing, TFS deficits have been observed in the detection of interaural time or phase differences via lateralization (e.g., Hawkins and Wightman, 1980; Häusler et al., 1983; Smoski and Trahiotis, 1986; Gabriel et al., 1992; Koehnke et al., 1995; Lacher-Fougère and Demany, 2005). Also studies on binaural masked detection or masking level differences (MLDs) have reported deficits in HI listeners (e.g., Hall et al., 1984; Staffel et al., 1990; Gabriel et al., 1992). In both tasks, lateralization and binaural detection, the interaural phase or time differences in the stimuli are coded in terms of phase-locking-based TFS processing (see Stern and Trahiotis, 1995, and Colburn, 1996). Apart from these *binaural* measures of TFS processing, frequency discrimination of tones with frequencies of up to 4-5 kHz is thought to be determined by a temporal mechanism based on phase locking (see Moore, 2003). Hence, deficits observed in the frequency discrimination of steady pure tones (e.g., Turner and Nelson, 1982; Tyler et al., 1983; Turner, 1987; Freyman and Nelson, 1991) and in the detection of low-rate FM (e.g., Zurek and Formby, 1981; Grant, 1987; Lacher-Fougère and Demany, 1998; Moore and Skrodzka, 2002; Buss et al., 2004) have been interpreted to indicate deficits in *monaural* TFS processing in HI listeners. This conclusion has been further supported by studies of frequency discrimination with harmonic complex tones (e.g., Horst, 1987; Moore et al., 2006; Hopkins and Moore, 2007). However, since none of the above mentioned studies has obtained both monaural and binaural measures of TFS processing, it remained unclear to what extent the deficits observed in the binaural tasks were due to monaural or independent binaural deficits.

Only a few studies have assessed the relation between TFS deficits and speech reception performance. Tyler et al. (1983), Glasberg and Moore (1989), Noordhoek

et al. (2001), and Buss et al. (2004) found significant correlations between frequency discrimination performance and word recognition in speech-shaped noise as well as quiet, while Horst (1987) did not find such correlations. Lorenzi et al. (2006) and Hopkins et al. (2008), using processed speech stimuli, presented evidence that HI listeners were less able to make use of the TFS information in speech than normal-hearing (NH) listeners. However, in these studies, the potential contribution of reduced frequency selectivity to the observed TFS deficits remained unclear. Reduced frequency selectivity might have affected the processing of TFS information in several ways (see also Moore, 2008a). For wideband signals, the outputs of broadened auditory filters would exhibit a more complex TFS than the outputs of “normal” filters (Rosen, 1987). In addition, the signal-to-noise ratio (SNR) in the presence of a wideband interferer would be smaller in the case of broadened filters, providing a less favorable input to the subsequent processing stages. Finally, parts of the preserved TFS information in the speech stimuli of Lorenzi et al. might have been coded in terms of FM-to-AM conversion through cochlear filtering (e.g., Zeng et al., 2004; Gilbert and Lorenzi, 2006). In such a case, filter broadening would result in reduced AM depths at the filter output and a less distinct representation of frequency transitions (e.g., downward and upward glides) across adjacent filters. Hence, the observed deficits in the TFS processing of wideband stimuli could, in principle, have resulted from reduced frequency selectivity rather than from deficits in subsequent auditory processing stages.

Therefore, the present study investigated potential deficits in phase-locking-based TFS processing, where possible effects of frequency selectivity should play a minor role. Nevertheless, the relation between frequency selectivity and TFS processing was examined here since both might be affected by a common underlying impairment factor such as outer hair cell (OHC) damage. The TFS processing was addressed binaurally through measurements of binaural masked detection and lateralization of pure tones with ongoing interaural phase differences (IPDs). As a complementary monaural measure, detection thresholds for low-rate frequency modulation (FMDTs) were obtained. The IPD thresholds and FMDTs were measured in quiet as well as in continuous noise backgrounds in order to test the robustness of the TFS processing to interfering noise. Physiological animal studies (e.g., Rhode et al., 1978; Abbas, 1981; Costalupes, 1985) have shown that phase locking to tones in the presence of

## 10 2. Relations between frequency selectivity, TFS processing, and speech reception

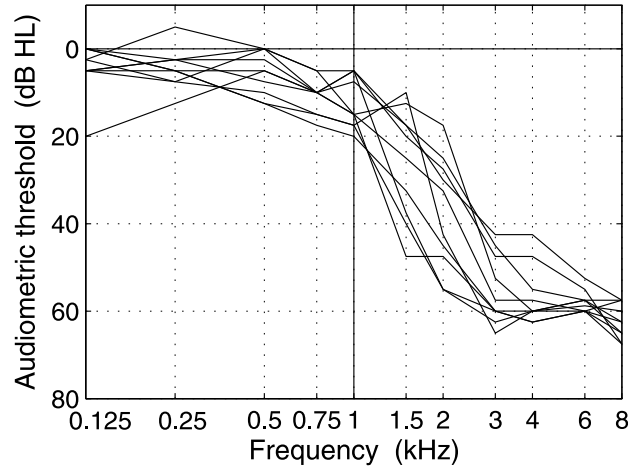


Figure 2.1: Audiograms of the ten HI listeners. For each listener, the mean of left and right ears is shown.

background noise is generally preserved at SNRs near behavioral detection thresholds but ceases at sufficiently low SNRs. However, as no comparable studies exist in impaired hearing, it cannot be excluded that hearing impairment might potentiate the susceptibility of phase locking to noise disturbance.

## 2.2 Methods

### 2.2.1 Listeners

The six NH listeners (three females and three males) were aged between 21 and 55 years (median: 28) and had audiometric thresholds better than 20 dB HL (ISO 389-8, 2004) at all octave frequencies from 125 to 8000 Hz and 750 to 6000 Hz. The ten HI listeners (three females and seven males) were aged between 24 and 74 years (median: 63). Their audiograms are shown in Fig. 2.1, and more detailed audiometric information is given in Table 2.1. Throughout the study, the HI subjects are sorted by age and the notation “HI<sub>*n*</sub>” is used to refer to the individual subject with index *n*. The audiograms were “normal” up to 1 kHz (thresholds  $\leq 20$  dB HL)

## 2.2 Methods

11

Table 2.1: Audiometric information for the ten HI listeners and the two listeners with OD. The ears that were tested on monaural FM detection are marked by asterisks.

ID	Gender	Age	Ear	Audiometric thresholds (dB HL)											Etiology
				125	250	500	750	1000	1500	2000	3000	4000	6000	8000	
HI <sub>1</sub>	F	24	L* R	5 0	-5 -5	0 0	5 5	15 15	25 25	35 30	60 55	60 55	55 65	60 70	Hypoxia at birth
HI <sub>2</sub>	M	53	L R*	5 5	0 5	5 0	10 10	5 5	15 20	30 30	45 40	40 45	50 55	55 60	Unknown
HI <sub>3</sub>	M	55	L* R	0 0	5 5	10 15	15 15	15 20	10 10	55 30 <sup>a</sup>	70 60	55 65	60 55	55 60	Noise induced
HI <sub>4</sub>	M	56	L* R	5 -5	5 0	5 -5	5 5	5 5	25 15	30 25	45 50	45 50	55 55	70 65	Hereditary
HI <sub>5</sub>	M	60	L* R	10 0	5 0	10 5	10 10	5 5	35 40	60 50	60 65	60 60	55 60	60 60	Noise induced
HI <sub>6</sub>	M	67	L* R	0 0	5 5	5 5	10 10	10 5	15 15	30 20	45 45	50 60	50 65 <sup>a</sup>	65 65	Unknown
HI <sub>7</sub>	M	70	L R*	5 5	5 10	0 5	10 10	15 15	10 15	20 15	50 55	65 55	60 60	60 65	Noise induced
HI <sub>8</sub>	F	70	L* R	5 5	5 5	15 10	20 15	20 20	30 35	45 45	60 60	65 60	60 60	60 55	Unknown
HI <sub>9</sub>	M	74	L R*	20 20	10 15	5 5	10 10	20 10	40 50	35 60 <sup>a</sup>	60 60	60 60	60 55	70 55 <sup>a</sup>	Noise induced
HI <sub>10</sub>	F	74	L* R	0 5	10 5	10 10	15 15	20 15	45 35	55 55	65 55	65 60	60 60	70 60	Noise induced
OD <sub>1</sub>	F	26	L R*	-5 -5	-5 -5	0 0	0 0	0 0	0 0	0 0	0 5	0 0	10 5	5 5	None
OD <sub>2</sub>	F	46	L R*	0 0	5 0	5 0	5 5	10 10	5 5	5 10	5 0	0 10	10 10	5 5	None

<sup>a</sup> Thresholds differ by more than 10 dB between the ears.

## 12 2. Relations between frequency selectivity, TFS processing, and speech reception

and sloping at higher frequencies to values of up to 70 dB HL. All listeners had bilaterally symmetric audiograms (within 10 dB, exceptions stated in Table 2.1), to avoid the issue of level balancing in binaural testing, as discussed in Durlach et al. (1981). The sensorineural origin of the hearing losses was established by means of bone-conduction measurements, tympanometry, and otoscopy. The etiologies stated in Table 2.1 were based on the subjects' reports. They ranged from hypoxia at birth (oxygen deficiency) and hereditary losses to noise-induced losses, either sudden or due to sustained exposure to intense sounds. The remaining two subjects had OD: Despite audiometric thresholds better than 15 dB HL at all test frequencies (see Table 2.1), they approached the research center, complaining about difficulties with understanding speech in noisy backgrounds. Their middle-ear status was normal and they did not report any history of otitis media or excessive noise exposure. ABRs were measured for these two subjects and the HI subject HI<sub>10</sub> (since HI<sub>10</sub> showed diverging results in the lateralization task). As the responses were normal, there was no indication of eighth-nerve tumors, brainstem lesions, or auditory neuropathy. Additionally, all listeners were screened on a binaural pitch task, testing the ability to hear a Huggins' pitch C-scale (Santurette and Dau, 2007). Santurette and Dau suggested that the absence of a binaural pitch percept might indicate the presence of a severe central auditory deficiency. Since all listeners in the present study perceived the pitch there was no indication of such a deficiency.

Each subject completed all tests, with the exception of one NH listener, for whom SRTs were not measured. The average testing time was 24 h per listener. All experiments were approved by the ethics committee of Copenhagen county.

### **2.2.2 Apparatus**

All stimuli were generated in MATLAB® and converted to analog signals using a 24-bit digital-to-analog converter (RME DIGI96/8). The sampling rate was 44.1 kHz for the speech reception measurement, 48 kHz for the masking and FM experiments, and 96 kHz for the lateralization task. The stimuli were presented in a double-walled sound-attenuating booth via Sennheiser HD580 headphones. Calibrations were done using a Brüel & Kjær (B&K) artificial ear (4153) and, prior to playing, 128-tap linear-

phase FIR equalization filters were applied to all broadband stimuli, rendering the headphone frequency response flat.

### 2.2.3 Statistical analyses

To accommodate the repeated-measures design, the statistical analyses were carried out using linear mixed-effects models (MEMs; Laird and Ware, 1982; Pinheiro and Bates, 2000), as implemented in S-PLUS®. The between-subject variability that was not explained in terms of the fixed effect subject group (or interactions of other fixed effects such as stimulus condition with subject group) was accounted for in terms of subject-specific random effects. In addition to analyses of variance (ANOVAs) and multiple comparisons of the fixed effects (with simultaneous 95% confidence intervals, either based on the Dunnett method or Monte Carlo simulations), the estimated random effects were extracted. They served as ranks for the individual listeners' performance on a given test, for example binaural masked detection or lateralization. In the following, the abbreviations SD and CI will be used for standard deviation and confidence interval, respectively.

## 2.3 Speech reception

### 2.3.1 Method

SRTs were measured for Danish closed-set Hagerman sentences (Dantale II, Wagener et al., 2003) in the presence of different interferers: a stationary speech-shaped noise (SSN) with the long-term spectrum of the Dantale II sentences, a sinusoidally and a randomly amplitude-modulated noise (SAM and RAM), a multitalker and a reversed two-talker background (MULTITALK and TWOTALK), and a dichotic, lateralized SSN (LATSSN). Specifically, the SAM noise was fully sinusoidally amplitude-modulated SSN, with a modulation rate of 8 Hz (cf. Füllgrabe et al., 2006). The RAM noise was randomly amplitude-modulated SSN, with the Hilbert envelope of a 20-Hz-wide noise used as modulator. The MULTITALK noise was a reversed 20-talker babble (supplied as track 3 on compact disc CD101R3 “Auditory Tests revised” by AUDiTEC of St. Louis). The TWOTALK noise consisted of running female and

## 14 2. Relations between frequency selectivity, TFS processing, and speech reception

male speech, with silent gaps longer than 250 ms removed, mixed at equal level, and time-reversed (supplied as tracks 8 and 9 on compact disc CD B&O 101 “Music for Archimedes” by Bang & Olufsen). The LATSSN (noise) was SSN which was lateralized to one side by means of a constant interaural time difference (ITD) of 740  $\mu$ s. For a given run, either the left or the right ear was leading, but the SRT was averaged across runs with lateralization to the left and right. In addition to these conditions with full-spectrum speech, two conditions with filtered speech were used,  $SSN_{\text{filt}}$  and  $SAM_{\text{filt}}$ , in which both target speech and interferer were lowpass filtered at 1 kHz (1024-tap FIR lowpass filter designed using the Parks-McClellan algorithm in MATLAB®). This was done to test the processing of speech information in the regions of normal hearing (as all listeners had normal audiometric thresholds up to 1 kHz). The SRTs in all the aforementioned conditions were measured binaurally with the target speech and interferer presented diotically, with the exception of the LATSSN condition, where the interferer was presented dichotically. In addition, SRTs in the SSN and SAM conditions were measured monaurally, for comparison with the other monaural tests of frequency selectivity and FM detection.

The SRT was defined as the SNR leading to 50% correct identification of the individual words in the Dantale II sentences. The interferer level was kept constant at 65 dB SPL while the sentence level was varied adaptively. In each condition, the listeners were trained on a single run of 20 sentences. Subsequently, the SRT was estimated as the average over two to three runs, depending on the condition. A monotonic improvement of threshold in a sequence of three runs was interpreted as a training effect. When such an effect occurred, further runs were taken until stable performance was reached, and the first runs were discarded. This procedure for dealing with training effects was applied to all the other tests in this study.

### 2.3.2 Results and discussion

Figure 2.2 shows the binaural SRTs for the NH (circles), the OD (bold numbers), and the HI listeners (plain numbers). The horizontal black bars denote the mean SRTs for the NH and HI listeners and the corresponding boxes represent  $\pm 1$  SD. Considering the first six conditions with full-spectrum speech, all listeners showed the lowest SRTs with the SAM and LATSSN interferers, while the highest SRTs were obtained



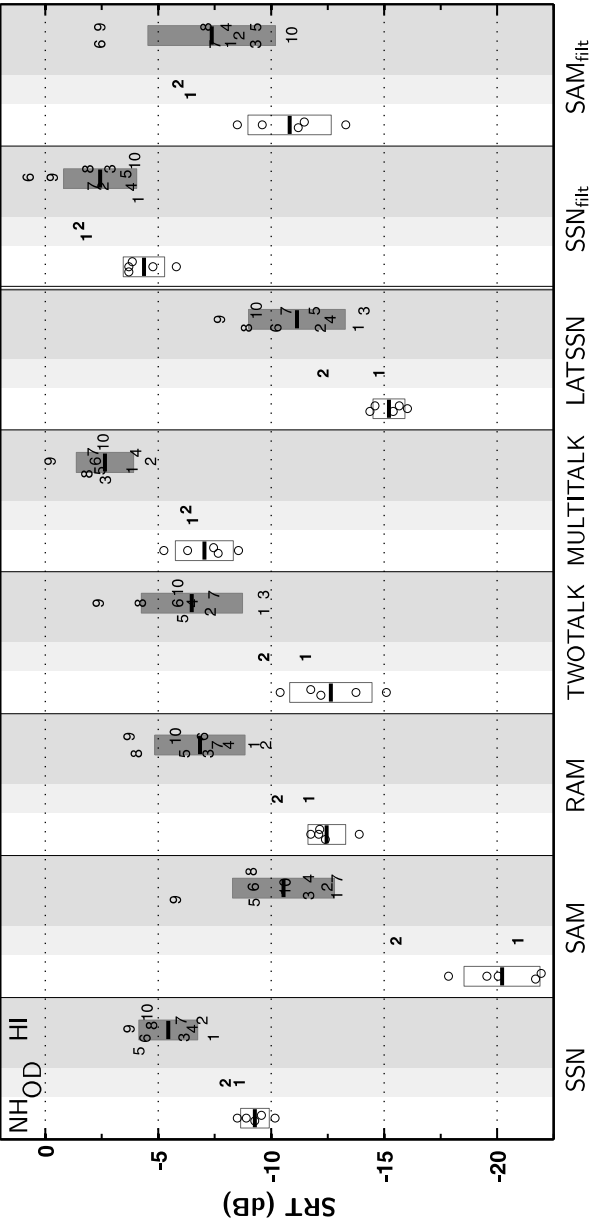


Figure 2.2: SRTs for the NH listeners (circles), the two listeners with OD (bold numbers), and the HI listeners (plain numbers). The different conditions are indicated at the bottom of each panel. The horizontal black bars show the mean SRTs for the NH and HI listeners and the corresponding boxes represent  $\pm 1$  SD.

## 16 2. Relations between frequency selectivity, TFS processing, and speech reception

with the MULTITALK interferer. The SRTs for the RAM and TWOTALK conditions lay slightly below those for the stationary SSN interferer. Performance in the conditions with lowpass-filtered speech ( $SSN_{\text{filt}}$  and  $SAM_{\text{filt}}$ ) was generally poorer than performance in the corresponding conditions with full-spectrum speech. An ANOVA was performed on the SRTs of the NH and HI listeners. The SRTs were found to be significantly higher for the HI listeners than for the NH listeners [ $F(1, 13) = 36.1$ ,  $p < 0.0001$ ]. The SRTs differed significantly across conditions [ $F(7, 91) = 238.7$ ,  $p < 0.0001$ ] and the interaction between listener group and condition was significant [ $F(7, 91) = 20.5$ ,  $p < 0.0001$ ]. Multiple comparisons revealed that the HI listeners performed more poorly than the NH listeners for all full-spectrum conditions [ $p < 0.001$ ]. For the two conditions with lowpass-filtered speech, the HI listeners' deficits were less pronounced. The deficit was significant for the  $SAM_{\text{filt}}$  condition [ $p < 0.01$ ], but not for the  $SSN_{\text{filt}}$  condition [ $p > 0.05$ ]. Within the group of HI listeners, no significant correlation was observed between the SRTs for the filtered speech and the full-spectrum speech [ $p > 0.05$ ]. Hence, they did not seem to make equally good use of the low-frequency and high-frequency information in the speech stimuli. For example, listeners  $HI_5$  and  $HI_{10}$  performed relatively well in the filtered-speech task, but poorly in the full-spectrum speech task.

Previously, Horwitz et al. (2002) measured speech reception performance of HI listeners in regions of normal hearing, using lowpass-filtered speech in an SSN masker. In contrast to the present results ( $SSN_{\text{filt}}$  condition), they found significantly poorer performance for their HI than for their NH listeners. However, their speech stimuli were presented at a level of 77 dB SPL, where a substantial spread of excitation on the basilar membrane (BM) would be expected, particularly toward places corresponding to higher characteristic frequencies (CFs). Consequently, they interpreted their finding in terms of a reduced ability of their HI listeners to encode the information at places with high CFs, where a hearing loss was present.

In addition to the SRTs, speech masking release was considered, i.e., the gain in terms of SRT for the SAM, RAM, TWOTALK, and LATSSN conditions when compared with the SRT for the stationary, diotic SSN condition. The group masking release values can be extracted from Fig. 2.2 as the differences in SRT between the corresponding conditions. The masking release values were significantly smaller for

### 2.3 Speech reception

17

the HI listeners than for the NH listeners [ $F(1, 13) = 21.8, p = 0.0004$ ]. While the SAM masking release values for the full-spectrum speech differed strongly between the NH and HI listeners [by 5.8 dB,  $p < 0.001$ ], the difference for the filtered speech just reached significance [1.5 dB,  $p = 0.05$ ]. The finding of less pronounced deficits with lowpass-filtered speech may, at least partly, be attributed to the fact that the HI listeners had normal low-frequency hearing thresholds and that the full-spectrum speech stimuli were not amplified to fully restore audibility at high frequencies. It is interesting that the HI listeners did not benefit from high-frequency information in terms of the SAM masking release: While the NH listeners showed a significantly larger masking release with full-spectrum speech (SAM–SSN) than with filtered speech (SAM<sub>filt</sub>–SSN<sub>filt</sub>) [difference in dB: 4.5 (CI 3.3,5.7)], the difference was not significant for the HI listeners [0.2 (CI –0.9,1.2) dB].

As mentioned above, higher SRTs were observed in the MULTITALK masker than in the SSN masker, independent of listener group. Hence, in addition to the energetic masking present for the latter, another detrimental masking effect must have limited speech intelligibility in the case of the MULTITALK background. This could, for example, have been the complex harmonic structure of the background babble, which interfered with the use of spectro-temporal cues in the target speech, such as formant transitions.

As can be seen in Fig. 2.2, the OD listeners showed rather small deficits in the reception of full-spectrum speech. Consistent with previous reports in literature (e.g., Middelweerd et al., 1990; Saunders and Haggard, 1992), they often performed at the lower limit of the NH group. Subject OD<sub>1</sub> showed elevated SRTs only in the two filtered-speech conditions. Subject OD<sub>2</sub> showed poorer performance than the NH listeners in all conditions except MULTITALK. Particularly in the SAM, LATSSN, and both filtered-speech conditions, her SRTs were increased relative to those for the NH group. Hence, for these two listeners, a deficit in speech reception was most apparent for the lowpass-filtered speech. This deficit might reflect a general difficulty understanding speech that is less redundant than full-spectrum speech (cf. Oxenham and Simonson, 2009). However, it could also reflect a specific problem with the processing of low-frequency information.

The monaural SRTs, which were measured only in the SSN and SAM condi-

## 18 2. Relations between frequency selectivity, TFS processing, and speech reception

tions, closely followed the corresponding binaural results described above (monaural SRTs were, on average, 1.5 dB higher than binaural SRTs). The mean monaural SSN SRT of  $-8.7$  (SD  $0.9$ ) dB for the NH listeners was consistent with the SRT of  $-8.4$  (SD  $1.0$ ) dB reported by Wagener et al. (2003). Since the monaural results do not provide any further insights they are not presented in detail.

## 2.4 Frequency selectivity

### 2.4.1 Method

Auditory filter shapes at 750 Hz were determined separately for each ear using a notched-noise paradigm (cf. Patterson and Nimmo-Smith, 1980). Rosen et al. (1998) presented evidence that auditory-filter shapes are output driven. Under the assumption of the power-spectrum model (cf. Patterson and Moore, 1986) that a constant SNR at the output of the auditory filter is required for detection, this is equivalent to saying that the filter shape is determined by the level of the target signal rather than the noise masker. Therefore, here, in order to obtain a faithful filter estimate, the signal level was kept constant while the masker level was varied adaptively. The 750-Hz target tones of 440-ms duration were presented at a fixed level of 50 dB SPL and were temporally centered in the 550-ms noise maskers. Maskers and tones were gated with 50-ms raised-cosine ramps. The noise was generated in the spectral domain as fixed-amplitude random-phase noise (this holds also for the noises in all remaining tests). Five symmetric ( $\delta f/f_0$ : 0.0, 0.1, 0.2, 0.3, and 0.4) and two asymmetric notch conditions ( $\delta f/f_0$ : 0.2|0.4 and 0.4|0.2) were used, where  $\delta f$  denotes the spacing between the inner noise edges and the signal frequency  $f_0$ . The outside edges of the noise maskers were fixed at  $\pm 0.8f_0$ .

A three-interval, three-alternative, forced-choice (3I-3AFC) weighted up-down method (Kaernbach, 1991) was applied to track the 75%-correct point on the psychometric function. A run was terminated after 14 reversals. The threshold was defined as the arithmetic mean of all masker levels following the fourth reversal. Following a training run for each notch condition, the threshold was estimated as the average over

three runs. If the SD of these three runs exceeded 1 dB, one or two additional runs were taken and the average of all was used.

A nonlinear minimization routine was implemented in MATLAB® to find the best-fitting rounded-exponential (roex) filter (e.g., Patterson and Moore, 1986) in the least-squares sense, assuming that the signal was detected using the filter with the maximum SNR at its output. Middle-ear filtering was taken into account, using the middle-ear transfer function supplied by Moore et al. (1997). However, the results presented in the following do not depend on this choice. Furthermore, besides the equivalent rectangular bandwidth (ERB) as a measure of filter tuning, also the 3-dB and 10-dB bandwidths were considered. However, because they yielded essentially identical results, for ease of comparison only the ERB results will be discussed further.

### 2.4.2 Results and discussion

The  $\text{roex}(p, r)$  filter model (Patterson et al., 1982) provided a good description of the individual notched-noise threshold data, with a residual root-mean-square (rms) fitting error of 0.64 (SD 0.25) dB, averaged across all subjects. Figure 2.3(a) shows the estimated ERBs for the NH and HI listeners as well as the two OD listeners. The HI listeners showed, on average, significantly higher bandwidths than the NH listeners [ $F(1, 14) = 13.5, p = 0.003$ ], by a factor of 1.2. However, the results varied considerably across the HI listeners, with four of them showing bandwidths in both ears within the range of the NH listeners. In addition to the ERB, significantly shallower lower and upper filter skirts were observed for the HI listeners than for the NH listeners [lower skirt:  $F(1, 14) = 10.9, p = 0.005$ ; upper skirt:  $F(1, 14) = 5.6, p = 0.03$ ].

As can be seen in Fig. 2.3(a), abnormal filter bandwidths were also found for the two OD listeners. While OD<sub>1</sub> showed significantly elevated bandwidths (compared to the NH group) in both ears, OD<sub>2</sub> showed an increased bandwidth only in the left ear. The difference between the ERB of the left and right ear (divided by the mean ERB of the two ears) is depicted in Fig. 2.3(b). While this interaural bandwidth asymmetry did not differ significantly between the NH and HI listeners, both OD subjects showed larger differences between the ears than the NH listeners and most of the HI listeners. Decreased frequency selectivity, as found here, has been reported previously in the OD literature (e.g., Narula and Mason, 1988; Saunders and Haggard, 1992). It is also

20 2. Relations between frequency selectivity, TFS processing, and speech reception

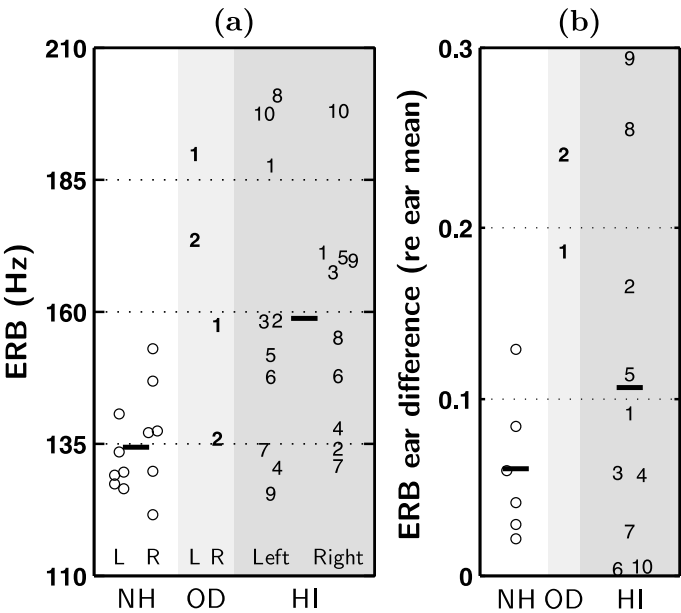


Figure 2.3: (a) ERB of the  $\text{roex}(p, r)$  filter estimates at 750 Hz for the NH listeners (circles), the two listeners with OD (bold numbers), and the HI listeners (plain numbers). For each group, the left and right symbols or numbers correspond to the left and right ears, respectively. The horizontal black bars denote group means. (b) Absolute value of the ERB differences between the ears, divided by the mean ERB for the two ears.

consistent with the finding of reduced distortion-product otoacoustic emission (OAE) amplitudes (Zhao and Stephens, 2006), if these are taken as an indication of OHC integrity.

The mean ERB of 134 (SD 9) Hz for the NH listeners is larger than the value of 106 Hz predicted by the ERB function given in Glasberg and Moore (1990). However, that function was designed to predict tuning in the presence of a masker with a constant spectrum level of about 35 dB SPL. It is known that the auditory filter bandwidth increases with increasing (output) level (e.g., Rosen et al., 1998). Hence, the larger bandwidths found here may be attributed to the higher masker levels applied (the average spectrum level here was 48 dB SPL). In fact, they are in good agreement with

the bandwidths reported by Moore et al. (1990) who measured at comparable masker levels.

## 2.5 Binaural masked detection

### 2.5.1 Method

The binaural masked thresholds for 750-Hz tones at fixed levels of 65 and 35 dB SPL were measured in bandlimited noise (50–1500 Hz). Three different masking conditions were tested: a diotic tone presented in a diotic noise ( $N_0S_0$ ), a diotic tone presented in an uncorrelated noise ( $N_uS_0$ ), and a tone with an interaural phase shift of  $180^\circ$  presented in a diotic noise ( $N_0S_\pi$ ). The first two conditions were measured using both tone levels whereas the last condition was measured only for the lower tone level. The tones of 500-ms duration were temporally centered in the 700-ms noise maskers. Maskers and tones were gated with raised-cosine ramps of 100-ms and 200-ms duration, respectively.

The same 3I-3AFC method as for the frequency selectivity measurement (including threshold estimation) was used. Also here, the signal level was kept constant while the masker level was varied adaptively. The final standard error of the masked threshold estimate, averaged across all listeners and conditions, was 0.4 dB.

### 2.5.2 Results and discussion

The masked thresholds for the NH, the OD, and the HI listeners are shown in Fig. 2.4, with SNRs given relative to the masker spectrum level. For all listeners, the thresholds were lower in the dichotic  $N_uS_0$  and  $N_0S_\pi$  conditions than in the corresponding diotic  $N_0S_0$  conditions. These MLDs reflected a release from masking in the dichotic configurations and will be discussed further below. An ANOVA revealed that the masked thresholds were significantly higher for the HI than for the NH listeners [ $F(1, 14) = 14.7$ ,  $p = 0.002$ ]. Furthermore, the masked thresholds differed significantly between the different binaural conditions [ $F(2, 59) = 536.9$ ,  $p < 0.0001$ ] and also the interaction between listener group and masking condition was significant [ $F(2, 59) = 4.2$ ,  $p = 0.02$ ]. While there was no significant

## 22 2. Relations between frequency selectivity, TFS processing, and speech reception

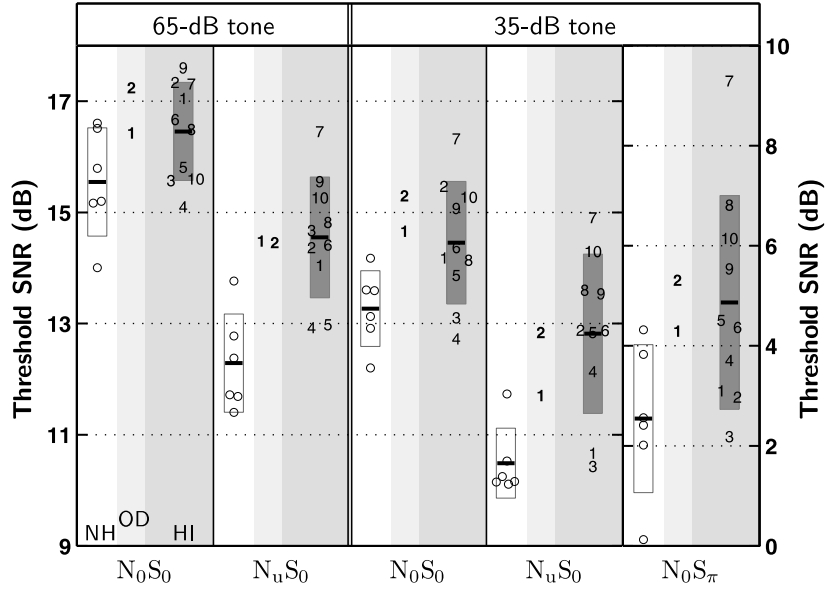


Figure 2.4: Binaural masked thresholds, i.e., tone level re masker spectrum level at detection threshold, for the NH listeners (circles), the two listeners with OD (bold numbers), and the HI listeners (plain numbers), obtained in three different masking conditions ( $N_0S_0$ ,  $N_uS_0$ , and  $N_0S_\pi$ ) and at two different tone levels (65 and 35 dB SPL). Note the different offset of the ordinate for the  $N_0S_\pi$  condition. Otherwise as Fig. 2.2.

difference between the NH and HI listeners for the diotic  $N_0S_0$  condition [group difference: 1.1 (CI  $-0.3, 2.4$ ) dB], thresholds for the dichotic conditions differed significantly [ $N_uS_0$  group difference: 2.3 (CI 1.0, 3.6) dB;  $N_0S_\pi$  group difference: 2.3 (CI 0.8, 3.9) dB]. Furthermore, within the group of HI listeners, a significant correlation between the  $N_uS_0$  and  $N_0S_\pi$  thresholds was observed [ $r = 0.87$ ,  $p = 0.001$ ]. Together, this suggests a deficit in TFS processing at threshold, which impaired  $N_uS_0$  and  $N_0S_\pi$  detection in similar ways.

Significantly larger SNRs were required for the detection of the 65-dB tones than for the 35-dB tones [effect of level on SNR: 1.9 (CI 1.5, 2.4) dB]. This is consistent with the notion of decreasing sharpness of the auditory filters with increasing tone level, if detector efficiency is assumed to be invariant (as found by Rosen et al., 1998). However, as can also be seen in Fig. 2.4 (mean results), the effect of tone level did



not differ significantly across masking condition or listener group. The latter is in agreement with Baker and Rosen (2002), who found a differential effect of tone level on the ERBs of their NH and HI listeners only for levels above 70 dB SPL.

The following MLDs were observed for the NH listeners:  $N_0S_0 - N_uS_0$  3.0 (SD 0.7) dB and  $N_0S_0 - N_0S_\pi$  10.7 (SD 1.3) dB. Since tone level had no significant effect, here, the  $N_0S_0 - N_uS_0$  MLD was averaged across the two tone levels. The HI listeners showed significantly smaller  $N_0S_0 - N_uS_0$  MLDs than the NH listeners [reduced by 1.3 (CI 0.1,2.4) dB]. However, no significant difference was found for the  $N_0S_0 - N_0S_\pi$  MLD [reduced by 1.1 (CI -0.2,2.5) dB]. Hence, the deficits in terms of the MLDs were less significant than the deficits in terms of the masked thresholds. This was due to the fact that the HI listeners exhibited not only significantly increased dichotic thresholds, but also slightly increased diotic thresholds, as previously reported by Staffel et al. (1990) and Gabriel et al. (1992).

Figure 2.4 also shows the masked threshold results for the two OD listeners. While subject OD<sub>2</sub> performed clearly more poorly than the NH listeners, subject OD<sub>1</sub> showed performance at the “lower edge” of that for the NH group. However, this applied to both the diotic and the dichotic masking conditions, as reported previously by Saunders and Haggard (1992). Therefore, in terms of their MLDs, no deficits were found for the OD listeners.

## 2.6 Lateralization

### 2.6.1 Method

Lateralization thresholds were measured for 750-Hz tones of 500-ms duration, at fixed levels of 70 and 35 dB SPL. The tones were gated synchronously and were lateralized by introducing a carrier-phase delay to one of the ears, giving rise to an IPD. For NH listeners, interaural carrier delays have been shown to dominate interaural gating delays for frequencies below about 1.5 kHz (see Zurek, 1993). To further weaken potential gating cues to lateralization, long onset/offset ramps of 200 ms each were used. Pilot measurements confirmed that the lateralization was solely based on TFS cues, since no significant difference was found between the lateralization thresholds

## 24 2. Relations between frequency selectivity, TFS processing, and speech reception

for tones with a waveform delay and tones with a carrier delay only. At each tone level, in addition to the lateralization threshold in quiet, three conditions with different bandlimited noise interferers (50–1500 Hz) were measured: diotic noise at a low (dioticLo) and a high sound level (dioticHi), and dichotic noise at an intermediate level (dichotic). The noise level in each condition was chosen relative to the individual's masked threshold ( $N_0S_0$  or  $N_uS_0$ ) to make sure that lateralization performance was not limited by tone detection and to reduce effects of frequency selectivity. The actual noise levels were as follows: dioticHi: 10 dB below masked threshold, for both tone levels; dioticLo: 40 dB below masked threshold for the 70-dB tones and 25 dB below masked threshold for the 35-dB tones; and dichotic: 20 dB below masked threshold for the 70-dB tones and 15 dB below masked threshold for the 35-dB tones.

A 2I-2AFC weighted up-down method was used to track 75% correct lateralization. The first interval always contained the zero IPD reference tone while the second interval contained the tone which was randomly lateralized to the left or right side. Listeners were instructed to indicate the direction of motion. The IPD was tracked logarithmically and the maximum IPD was restricted to  $90^\circ$ , since the extent of lateralization starts to decline for values above  $90^\circ$  (Kunov and Abel, 1981). The background interferer was presented continuously during the whole run. A run was terminated after 14 reversals and the threshold was defined as the geometric mean of all IPD values following the fourth reversal. Listeners were trained in at least two sessions and performed more than 1200 lateralization judgements (constant stimuli) prior to actual data collection. IPD threshold was estimated as the geometric mean over three runs. If the SD over these runs, relative to the mean IPD threshold, exceeded a factor of 0.2 (which corresponds to a constant criterion in logarithmic units), additional runs were taken and the average of all was used. The final relative standard error of the IPD threshold estimate, averaged across all listeners and conditions, was 0.13.

### 2.6.2 Results and discussion

The analysis of the lateralization results was performed on the log-transformed IPDs, as these satisfied the requirements of normal error distributions. This is in line with previous reports in literature on lateralization (e.g., Saberi, 1995; Lacher-Fougère and

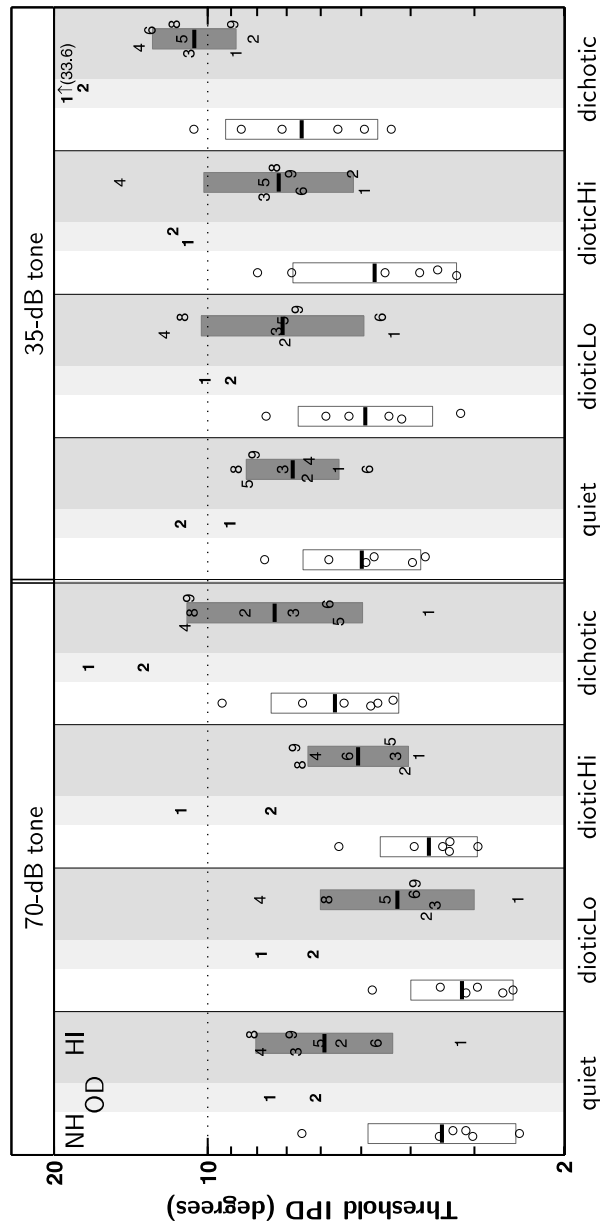


Figure 2.5: Lateralization thresholds for the NH listeners (circles), the two listeners with OD (bold numbers), and the HI listeners (plain numbers), at two different tone levels (70 and 35 dB SPL) and for different interferer conditions (see text). Otherwise as Fig. 2.2.

## 26 2. Relations between frequency selectivity, TFS processing, and speech reception

Demany, 2005). Figure 2.5 shows the IPD thresholds for the NH, OD, and HI listeners. The HI subjects HI<sub>7</sub> and HI<sub>10</sub> (not shown) performed much more poorly on lateralization than the remaining HI listeners. Therefore, their IPD thresholds were not included in the group averages and will be discussed separately further below. However, the conclusions presented in the following would remain unchanged if they were taken into account. Two trends can be seen in Fig. 2.5. First, lateralization performance was better at the higher tone level than at the lower level. Second, the HI listeners showed generally higher IPD thresholds than the NH listeners.

An ANOVA confirmed both the significant difference between NH and HI listeners [ $F(1, 12) = 8.7$ ,  $p = 0.01$ ], and the effect of tone level [ $F(1, 94) = 71.5$ ,  $p < 0.0001$ ]. The effect of interferer condition was also significant [ $F(3, 94) = 27.8$ ,  $p < 0.0001$ ], while interactions did not reach significance. The dichotic noise conditions led to the highest IPD thresholds, although the noise levels were actually lower than in the dioticHi conditions. This may, at least partly, be attributed to the fact that the dichotic noise gave rise to a diffuse, broad percept, while the diotic noise was lateralized in the midline. Hence, the latter provided an additional, ongoing reference cue since the noise was switched on continuously during a run.

Comparing performance in the dioticLo and dioticHi conditions, the HI listeners seemed to cope as well as the NH listeners with the increase of noise level. Generally, noise did not have a greater effect upon lateralization performance for the HI listeners than for the NH listeners, irrespective of tone and noise levels (as reflected in a lack of interaction between listener group and condition). Apart from higher thresholds for the HI listeners, the two groups of listeners showed a very similar pattern of results across conditions, with one exception, the quiet condition at the high tone level (left-most panel in Fig. 2.5). Here, the lateralization thresholds for the HI listeners were a factor of 1.7 higher than for the NH listeners, while in the other conditions thresholds were, on average, a factor of 1.4 higher. For the dichotic condition at the same tone level (factor of 1.3), one might have expected a larger deficit than in the quiet condition: While in both conditions an ongoing reference cue was absent, a smaller fraction of nerve impulses would have been expected to be phase locked to the tone in the presence of the noise interferer, thus possibly producing more difficulties for the impaired auditory system. This was, however, not the case. Also, the HI listeners'

deficit in quiet was actually smaller at the lower tone level (factor of 1.4) than at the higher level.

Hawkins and Wightman (1980) and Smoski and Trahiotis (1986) reported different effects of stimulus level on lateralization performance. For HI listeners with similar audiograms as in the present study, they measured lateralization thresholds in quiet at a low and a high stimulus level, in regions of normal hearing. For narrowband noise stimuli, the HI listeners MM and MD in Hawkins and Wightman (1980) showed a smaller lateralization deficit at the higher stimulus level than at the lower level. In contrast to this, and consistent with the present results, Smoski and Trahiotis (1986) observed a larger deficit in lateralization at the higher level using pure tones. In the same study, this trend was less clear when using narrowband noise stimuli. Hence, the discrepancy between the studies may be at least partly attributable to the differences in the stimuli.

Smoski and Trahiotis (1986) suggested that the lateralization judgement at high levels could be based on the excitation of a large portion of the BM rather than only on local excitation, and that a hearing loss might affect the integration of the non-local information. This interpretation is consistent with the present results for lateralization in quiet and in noise. At the tone level of 70 dB SPL, one would expect a substantial spread of excitation, particularly towards places that correspond to higher CFs. The NH listeners might have integrated the additional information present at these high-frequency places, whereas the HI listeners might not have been able to benefit from this information, as it fell in the sloping region of their hearing loss. Indeed, if actually included, information from defective neural units (as e.g., desynchronized information across frequencies) might have had a detrimental effect on lateralization acuity. The role of spread of excitation is reduced at the lower tone level of 35 dB, but also at the higher level of 70 dB in the presence of background noise, as the latter partly masks non-local excitation. This would explain why the deficit observed for the HI listeners (relative to NH) was largest at the high tone level in quiet.

As mentioned above, the HI subjects HI<sub>7</sub> and HI<sub>10</sub> performed more poorly on lateralization than the remaining HI listeners. Subject HI<sub>7</sub> showed markedly increased lateralization thresholds, independent of interferer condition and tone level. His IPD thresholds ranged from 21° to 27° at the high tone level, and 32° to 40° at

## 28 2. Relations between frequency selectivity, TFS processing, and speech reception

the low tone level, without showing a particular susceptibility to noise interference. For subject HI<sub>10</sub>, lateralization thresholds could not be determined. Even after a considerable amount of training, her performance remained at chance level (even at the maximum IPD of 90°).<sup>1</sup>

The two OD listeners showed markedly higher lateralization thresholds than the NH listeners, for all interferer conditions and at both tone levels (see Fig. 2.5). On average, the IPD thresholds for subjects OD<sub>1</sub> and OD<sub>2</sub> were increased relative to those for the NH listeners, by factors of 2.6 and 2.2, respectively. Both showed the most pronounced problems with lateralization in the presence of the dioticHi and dichotic noise interferers. In fact, in these conditions, they performed even more poorly than most of the HI listeners.

## 2.7 Frequency modulation detection

### 2.7.1 Method

Detection thresholds for sinusoidal frequency modulation (FMDTs) were measured monaurally for carrier frequencies of 125, 750, and 1500 Hz. Prior to gating, the stimulus was a frequency-modulated sinusoid defined by:

$$s(t) = a \sin \left[ 2\pi f_c t + \frac{\Delta f}{f_m} \sin(2\pi f_m t + \varphi) \right], \quad (2.1)$$

where  $f_c$  represents the carrier frequency,  $\Delta f$  the maximum frequency excursion, and  $f_m$  the FM rate. The FM phase  $\varphi$  was always  $1.5\pi$ . The phase-locking-based temporal mechanism for FM detection has been found to be operative only at FM rates below 10 Hz, whereas at higher rates, FM detection is thought to be based primarily on an FM-to-AM conversion mechanism (e.g., Moore and Sek, 1996; Lacher-Fougère and

<sup>1</sup> While subject HI<sub>7</sub> showed consistently poor performance on all TFS-processing tests (poorest performance of all listeners on binaural masked detection and FM detection), subject HI<sub>10</sub>, who was not able to lateralize at all, showed relatively poor performance on masked detection, but average performance in the FM detection task. Although it was ensured that HI<sub>10</sub> had understood the lateralization task, it cannot be excluded that her problem was, at least partly, due to the nature of the 2I-2AFC task, rather than a problem with lateralization per se.

## 2.7 Frequency modulation detection

29

Demany, 1998).<sup>2</sup> Here, both mechanisms were tested, by using FM rates of 2 Hz and 16 Hz. The tone levels were 30 dB sensation level (SL; individual hearing thresholds were determined by means of 3I-3AFC detection measurements) and 70 dB SPL. The impact of noise interference was tested by measuring the FMDT for 2-Hz FM tones at 750 Hz in a bandlimited noise (50–1500 Hz), at a level 10 dB below the individual masked threshold. At 1500 Hz, all measurements were undertaken in the presence of a low-level noise background (50–3000 Hz, with a spectrum level 55 dB below the tone level), in order to mask low-frequency cues due to spread of excitation.

Finally, in order to assess the phase-locking-based mechanism further, similar to the paradigm used by Moore and Sek (1996), FMDTs for 2-Hz FM tones with a superimposed AM were measured at the carrier frequencies of 750 and 1500 Hz. In view of the findings of Grant (1987), who observed a significantly larger deficit in FM detection in HI listeners if the FM tones were randomly rather than sinusoidally amplitude modulated, here, a quasi-sinusoidal AM was used: While the modulation depth was fixed at a peak-to-valley ratio of 6 dB, the instantaneous modulation rate either increased or decreased as a linear function of time. According to Moore and Sek (1996), the peak-to-valley ratio of 6 dB should be large enough to disrupt FM-to-AM conversion cues, but still small enough not to induce substantial level-related pitch shifts. Hence, for the conditions with added AM, the amplitude  $a$  in Equation (2.1) was time dependent,

$$a(t) \propto 1 + m \sin(2\pi F_a(t) + \vartheta). \quad (2.2)$$

Here,  $m$  represents the AM depth and  $F_a(t)$  is the integral of the instantaneous modulation rate

$$F_a(t) = \int_0^t d\tau \left( f_1 + \frac{f_2 - f_1}{T} \tau \right), \quad (2.3)$$

with  $T$  representing the tone duration. The initial and final modulation rates  $f_1$  and

<sup>2</sup> Chen and Zeng (2004) measured FMDTs in cochlear-implant subjects and provided further evidence for the temporal mechanism. Since only a single electrode was stimulated, FM-to-AM conversion (or excitation-pattern) cues were absent. Nevertheless, FM could be detected even for FM-rates as high as 320 Hz, indicating that the boundary for the temporal mechanism might be higher than 10 Hz. However, in acoustic hearing, the FM-to-AM conversion mechanism most likely determines FM detection performance at such high FM rates. This is consistent with the relatively poor performance of the cochlear-implant subjects: Their FMDTs were one to two orders of magnitude poorer than the ones of the NH listeners.

## 30 2. Relations between frequency selectivity, TFS processing, and speech reception

$f_2$  were each chosen randomly out of the interval between 1 and 3 Hz, under the constraint  $|f_2 - f_1| > 1$  Hz. Also the AM phase  $\vartheta$  was randomized. Independent of condition, the FM tones had a duration of 750 ms and were gated with 50-ms raised-cosine ramps.

A 3I-3AFC weighted up-down method was used to track 75% correct FM detection. In the conditions without AM, two of the intervals contained unmodulated tones, whereas the target interval contained the FM tone. In the conditions with added AM, all three intervals were independently amplitude modulated and the listeners were instructed to detect the interval containing the FM by listening for its characteristic high-low-high warble. The maximum frequency excursion  $\Delta f$  was tracked logarithmically. A run was terminated after 12 reversals and the threshold was defined as the geometric mean of all  $\Delta f$  values following the fourth reversal. Prior to data collection, a training session was given in which the listeners were trained on all conditions. Initially, both ears were tested on 2-Hz FM detection at 750 Hz in quiet and subsequently the worse ear was chosen for further testing. This was done in order to obtain the largest possible range of FMDTs among the HI listeners, particularly in view of the subsequent comparison with the results of the other tests such as frequency selectivity. Furthermore, it seemed reasonable to assume that the worse ear was limiting the binaural TFS-processing performance, particularly in the lateralization task. The FMDT was estimated as the geometric mean over three runs. If the SD over these runs, relative to the mean FMDT, exceeded a factor of 0.15, additional runs were taken and the average of all was used. The final relative standard error of the FMDT estimate, averaged across all listeners and conditions, was 0.08.

### 2.7.2 Results and discussion

The analysis of the FM detection results was performed on the log-transformed FMDTs, as these satisfied the requirements of normal error distributions. This is in agreement with previous reports in literature on FM detection (e.g., Zurek and Formby, 1981; Buss et al., 2004). For all listeners, FM detection performance did not differ significantly between the tone levels of 30 dB SL and 70 dB SPL [two-tailed t-test:  $p = 0.79$ ]. Therefore, only the 30-dB results are considered in the following. Figure 2.6 shows the FMDTs for the NH, OD, and HI listeners. As can be seen, for all



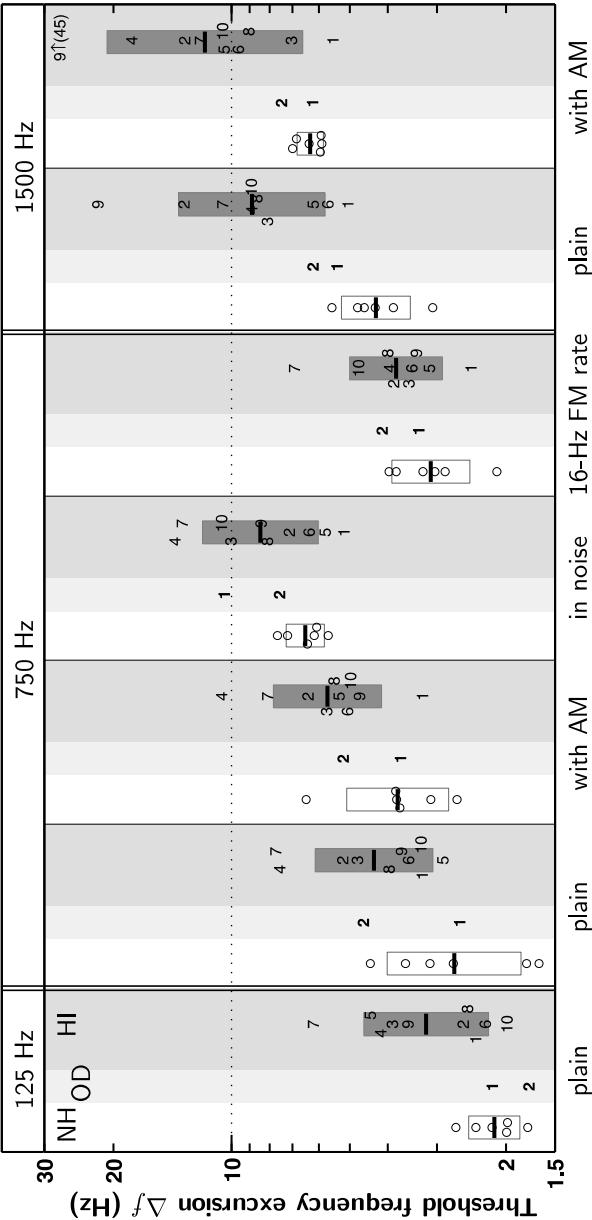


Figure 2.6: FMDTs for the NH listeners (circles), the two listeners with OD (bold numbers), and the HI listeners (plain numbers), for three different carrier frequencies (125, 750, and 1500 Hz) and for different measurement conditions (see text; “plain” refers to 2-Hz FM in quiet). All results were obtained at 30 dB SL. Otherwise as Fig. 2.2.

## 32 2. Relations between frequency selectivity, TFS processing, and speech reception

groups, the FMDTs increased with increasing carrier frequency, consistent with previous studies (e.g., Demany and Semal, 1989). The HI listeners performed generally more poorly than the NH listeners. On average, their FMDTs were a factor of 1.5 higher than for the NH listeners.

An ANOVA confirmed the statistical significance of the group difference [ $F(1, 14) = 16.9, p = 0.001$ ], as well as the effect of tone frequency [ $F(1, 89) = 56.7, p < 0.0001$ ]. No significant interaction between listener group and tone frequency was observed [ $p = 0.19$ ]. While the log-transformed FMDTs increased linearly as a function of frequency, the Weber fractions  $\Delta f / f_c$  decreased from 125 to 750 Hz by a factor of 4 and then remained constant up to 1500 Hz. Zurek and Formby (1981) measured FMDTs in HI listeners and found larger deficits for low-frequency tones than for high-frequency tones, given the same degree of hearing loss ( $< 30$  dB HL) at the test frequency. However, the FM detection deficits at 125 Hz observed in the present study were substantially smaller than the ones reported in that study. This might be due to the fact that the HI listeners of Zurek and Formby (1981) showed slightly higher audiometric thresholds at 125 Hz and generally more severe losses below 1000 Hz than the HI listeners of the present study.

FMDTs differed significantly across measurement conditions [2-Hz FM in quiet (“plain”), added AM, noise interference, and higher FM rate;  $F(3, 89) = 24.1, p < 0.0001$ ]. The interaction between listener group and measurement condition reached only marginal significance [ $F(3, 89) = 2.5, p = 0.07$ ]. However, for the following multiple comparison analysis, the interaction term was kept in the MEM.

As revealed by the multiple comparisons, the group differences between NH and HI listeners were significant for the 2-Hz FM in quiet and the condition with added AM [group difference in terms of  $\log_{10}(\text{FMDTs})$  for 2-Hz FM: 0.23 (CI 0.09, 0.37); group difference with added AM: 0.20 (CI 0.04, 0.35)]. For all listeners, the FMDTs with added AM were increased relative to those for the condition with FM only. However, as this increase was similar for the NH and HI listeners, it seems that both groups relied to a comparable extent on FM-to-AM conversion cues, when AM was absent. No significant group difference was found in the condition with the higher FM rate of 16 Hz [group difference: 0.09 (CI  $-0.08, 0.26$ )]. Thus, regarding the different FM rates (2 Hz vs 16 Hz), the HI listeners showed a significant deficit on FM de-

tection at the low rate but not at the high rate, where the FM-to-AM conversion is supposed to be the dominant detection mechanism. This can be seen in Fig. 2.6 (second and fifth panels): While the HI listeners’ performance was better for the higher FM rate, the NH listeners’ performance was worse. Taken together, this suggests that the observed deficits in the detection of 2-Hz FM were indeed due to problems with phase-locking-based TFS processing.

In the presence of the noise interferer, all listeners performed worse than in quiet. However, the HI listeners did not perform significantly more poorly than the NH listeners in this condition [group difference: 0.11 (CI  $-0.06, 0.29$ )]. Hence, the HI listeners did not show an increased susceptibility to noise interference. This is in agreement with the results of Turner (1987), who measured pure-tone frequency difference limens in the presence of low-frequency masking noise for four NH and four HI listeners and found a similar effect of the noise upon performance for the two groups of listeners. Also, Horst (1987) measured frequency discrimination in noise. However, the question of a different impact of noise on the performance of the NH and HI listeners could not be addressed, since he did not measure the frequency difference limen for a given noise level but determined the noise level at which a given fixed frequency difference could just be perceived.

Figure 2.6 also shows the FMDTs for the two OD listeners. Their FMDTs did not differ substantially from those for the NH listeners. Subject OD<sub>2</sub> performed at the “lower edge” of the NH listeners except for the 125-Hz carrier, where her performance was good. For subject OD<sub>1</sub>, a deficit was observed for the 750-Hz carrier with interfering noise. Otherwise her performance was essentially normal.

## 2.8 Comparison of results across tests

### 2.8.1 Hearing-impaired listeners

Pearson correlations and two-tailed  $p$  values were examined to study the relations between the results of the different auditory tests within the group of HI listeners. The findings are schematized in Fig. 2.7.

## 34 2. Relations between frequency selectivity, TFS processing, and speech reception

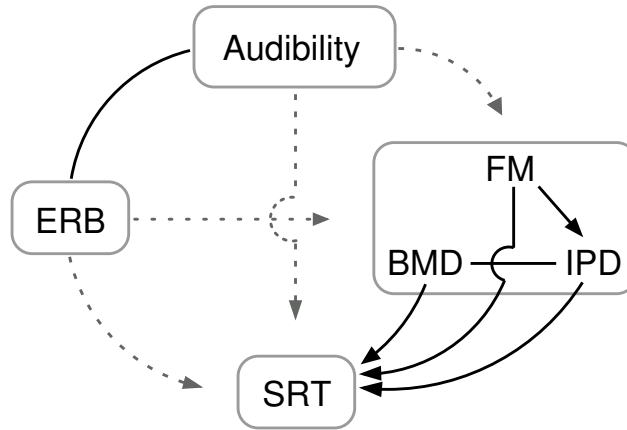


Figure 2.7: Relations between the results for the different auditory tests within the group of HI listeners: pure-tone hearing thresholds (audibility), frequency selectivity (ERB), monaural frequency-modulation detection (FM), binaural masked detection (BMD), tone lateralization (IPD), and speech reception (SRT). Solid lines indicate significant correlations whereas dotted lines indicate correlations that were not significant. The direction of the arrows is solely based on the assumed sequence of processing in the auditory pathway. Therefore, arrowheads were omitted where the order is uncertain or where the processing might take place in parallel rather than in sequence.

### *Correlations with absolute hearing thresholds*

Frequency selectivity in terms of the ERB at 750 Hz was significantly correlated with the individual hearing threshold at this frequency [ $r = 0.77$ ,  $p = 0.009$ ]. Here, the hearing threshold was estimated by means of a 3I-3AFC method with a 1-dB step-size. When the standard audiometric threshold (with a 5-dB stepsize) was considered instead, the correlation was smaller [ $r = 0.53$ ], but increased when thresholds were averaged in terms of the pure-tone average (PTA) threshold at 0.5, 1, 2, and 4 kHz [ $r = 0.8$ ]. The finding of a correlation between frequency selectivity and hearing threshold is consistent with previous reports in literature (e.g., Tyler et al., 1983; Moore, 1996), although less distinct correlations have been observed for hearing losses below 30 dB HL (see Baker and Rosen, 2002).

No significant correlations were observed between individual hearing thresholds and performance in the three tests of TFS processing (binaural masked detection, lateralization, and FM detection). Tones with equal sound pressure levels were used

for all listeners in the masked detection and lateralization tasks. Hence, the deficits in performance that were observed at the low tone level of 35 dB SPL could have been due to the slightly differing sensation levels (ranging from 32 to 38 dB SL for the NH group and 23 to 34 dB SL for the HI group). However, the absence of correlations between hearing thresholds, and thereby sensation levels, and masked detection/lateralization performance makes this unlikely. With regard to FM detection, subject HI<sub>9</sub>, who showed markedly worse performance at 1.5 kHz, also had the highest hearing thresholds at this frequency. Nevertheless, the correlation between the hearing thresholds and FMDTs at 1.5 kHz was not significant when considering all HI listeners [ $r = 0.39$ ,  $p = 0.27$ ]. Finally, the hearing thresholds were not significantly correlated with the results for speech reception, regardless of whether the hearing thresholds at single frequencies or averages across frequencies were considered.

The absence of correlations with the hearing thresholds can, to some extent, be attributed to the homogeneity of the HI group in terms of their audiograms. Also, given the limited number of listeners, only rather strong correlations would be expected to be significant. Hence, here and in the following, the absence of a significant correlation does not necessarily imply the absence of a relationship.

### ***Correlations between the various tests of TFS processing and frequency selectivity***

The deficits observed for the HI listeners with binaural masked detection, lateralization, and FM detection provide strong evidence for deficits in phase-locking-based TFS processing. However, no significant correlations were observed between frequency selectivity and these tests of TFS processing.<sup>3</sup> This can be illustrated by means of individual results among the HI listeners: Subject HI<sub>1</sub> showed poor frequency selectivity, but good TFS-processing skills, whereas subject HI<sub>7</sub> performed well on the former, but poorly on the latter. Subject HI<sub>10</sub> showed poor performance in both domains. Hence, it seems that the deficits found in TFS processing cannot be attributed

<sup>3</sup> At first, it may seem surprising that the diotic masked thresholds ( $N_0S_0$ ) and frequency selectivity were not correlated. However, in addition to the filter bandwidth, the masked thresholds are determined by the detector efficiency, i.e., the SNR at the output of the auditory filter required for detection.

## 36 2. Relations between frequency selectivity, TFS processing, and speech reception

solely to a deficit in frequency selectivity, but must be, at least partly, due to another impairment factor. This is further supported by the finding of TFS-processing deficits in quiet, which cannot be explained in terms of frequency selectivity.

Significant correlations were found among the tests of TFS processing. When correlations between the tests were observed for multiple test conditions, such as for the different interferer conditions in the lateralization task, an overall correlation is given in the following, instead of reporting the correlations for each individual condition. The overall correlation is based on the listeners' average performance on that test. This average performance was measured in terms of the estimated random effect, which summarizes individual performance across multiple conditions. Here, it represents the performance deviation of an individual HI listener from the HI group mean. Since the random effect accounts for multiple measurement conditions simultaneously, the corresponding correlation results are more robust and more conservative in terms of significance. Using this statistic, a significant correlation was observed between lateralization performance and the binaural masked thresholds in the  $N_0S_\pi$  condition [ $r = 0.80$ ,  $p = 0.01$ ], as has been observed previously (Hall et al., 1984; Kinkel et al., 1988; Koehnke et al., 1995). While the correlation between lateralization performance and the  $N_uS_0$  thresholds was rather marginal [ $p \sim 0.08$ ], no such correlation was observed for the  $N_0S_0$  thresholds [ $p > 0.2$ ].<sup>4</sup> The above correlation between lateralization performance and  $N_0S_\pi$  detection thresholds remained significant when controlling for individual hearing thresholds by means of partial correlation [ $r = 0.83$ ,  $p = 0.01$ ].

Performance on monaural FM detection and binaural masked detection was not correlated significantly.<sup>5</sup> However, the monaural FMDTs at 750 Hz were significantly correlated with lateralization performance [ $r = 0.79$ ,  $p = 0.01$ ]. Considering the different FM conditions separately, the correlations were strongest for the conditions with

<sup>4</sup> The fact that no significant correlation was observed between the IPD thresholds and the  $N_0S_0$  or  $N_uS_0$  masked thresholds is not surprising, as the levels of the diotic and dichotic noise interferers in the lateralization task had been chosen according to the individual  $N_0S_0$  and  $N_uS_0$  detection thresholds, in order to make sure that lateralization performance was not limited by tone detection.

<sup>5</sup> A reason for this could be that FM detection constituted a suprathreshold measure of TFS processing, while masked detection assessed the latter at threshold. Apart from this, since the tone detection could have been accomplished monaurally, it seems reasonable to assume that the binaural detection performance was not solely determined by the “worse” ear, which was tested on FM detection.

noise interference and with added AM. The correlation remained significant when controlling for individual hearing thresholds by means of partial correlation [ $r = 0.79$ ,  $p = 0.02$ ]. The fact that binaural and monaural (suprathreshold) TFS processing were correlated for the HI listeners suggests that the binaural deficit might be mainly attributable to a monaural impairment factor.

### *Correlations with speech reception*

As depicted in Fig. 2.8, two of the full-spectrum speech conditions, LATSSN and TWOTALK, showed significant correlations with the measures of TFS processing while no significant correlations were observed for the other speech conditions, including filtered speech.<sup>6</sup> Performance in the dichotic masked detection tasks (conditions  $N_0S_\pi$  and  $N_uS_0$ , in terms of the estimated random effects) was correlated with the SRTs in the LATSSN condition [ $r = 0.85$ ,  $p = 0.002$ ]; see Fig. 2.8(a). The correlation was also significant when the masking release instead of the SRT was considered [ $r = 0.80$ ,  $p = 0.005$ ]. For the sake of brevity in the following, a correlation with the masking release will only be given if it was stronger than the correlation with the corresponding SRT itself. The SRTs in the LATSSN condition were also significantly correlated with lateralization performance, but only for the dioticHi condition at the high tone level [ $r = 0.80$ ,  $p = 0.02$ ], see Fig. 2.8(b).<sup>7</sup> The pattern of correlations between the LATSSN SRTs and the masked thresholds as well as the lateralization thresholds remained unchanged when partialing out the individual

<sup>6</sup> In the filtered-speech conditions, listeners HI<sub>6</sub> and HI<sub>9</sub> performed markedly more poorly than all other listeners (Fig. 2.2). Subject HI<sub>9</sub> showed the largest deficits in speech reception among the HI listeners. However, his poor performance on FM detection at 1.5 kHz, which might have been a sign of substantial deficits in the processing of high-frequency information, cannot account for the deficits in the reception of lowpass-filtered speech. Similarly, subject HI<sub>6</sub>'s problems with lowpass-filtered speech were not reflected in his performance on the auditory tests of frequency selectivity or TFS processing. The reason for this remains unclear.

<sup>7</sup> Given that a dichotic noise interferer was used in the LATSSN speech condition, it might seem counterintuitive that a correlation was found in the case of the dioticHi lateralization condition but not the dichotic condition. However, the dichotic noise interferer (as compared to a diotic one) exerted rather opposite effects on speech reception and lateralization: While it gave rise to a release from masking in the speech task, it represented an additional challenge in the lateralization task. Furthermore, the level of the dioticHi noise in the lateralization task was comparable to the level of the noise interferer in the speech task (if the levels are considered relative to the corresponding masked thresholds for tone and speech, respectively).

## 38 2. Relations between frequency selectivity, TFS processing, and speech reception

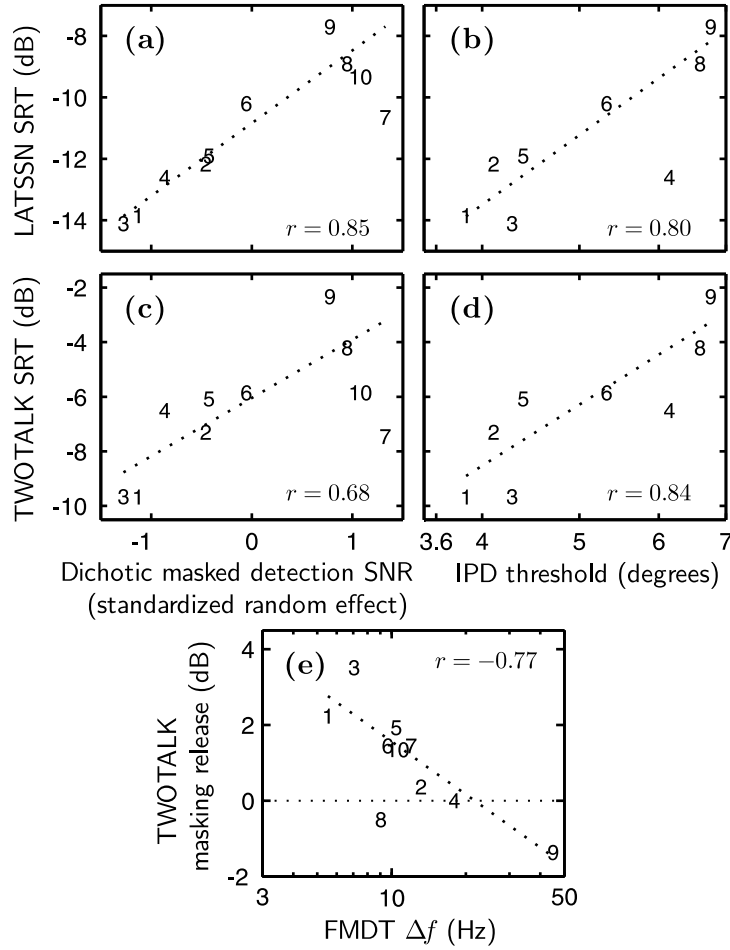


Figure 2.8: Correlations between performance on speech reception and TFS processing within the group of HI listeners. The dotted regression lines were obtained by means of least trimmed squares robust regression (Rousseeuw, 1984). (a) Correlation between the LATSSN SRTs and performance for dichotic masked detection ( $N_u S_0$  and  $N_0 S_\pi$  conditions). The latter is given in terms of the standardized random effects, which measure the individual deviations from the HI group mean. Better/worse than average performance, i.e., a smaller/larger threshold SNR, results in a negative/positive random effect. The interval from  $-1$  to  $1$  covers 68% of the HI “population”. (b) Correlation between the LATSSN SRTs and the IPD thresholds in the dioticHi condition for the 70-dB tones. (c/d) Same as (a/b) but for TWOTALK SRTs. (e) Correlation between the TWOTALK masking release (re SSN) and the FMDTs for 1.5-kHz tones with added AM.



## 2.8 Comparison of results across tests

39

hearing thresholds [masked detection:  $r = 0.82$ ,  $p = 0.007$ ; lateralization:  $r = 0.76$ ,  $p < 0.05$ ]. For the TWOTALK condition, significant correlations were found with both the dichotic masked thresholds ( $N_0S_\pi$  and  $N_uS_0$ ) and the lateralization thresholds in the dioticHi condition [masked detection:  $r = 0.68$ ,  $p = 0.03$ ; lateralization:  $r = 0.84$ ,  $p = 0.009$ ], as can be seen in Fig. 2.8(c) and (d), respectively. While the correlation with the masked thresholds was marginal when controlling for the individual hearing thresholds, the correlation with the lateralization thresholds remained significant [masked detection:  $r = 0.60$ ,  $p = 0.09$ ; lateralization:  $r = 0.81$ ,  $p = 0.03$ ].

No significant correlations between performance on speech reception and FMDTs at 125 and 750 Hz were found. However, at 1.5 kHz, the FMDTs with added AM were significantly correlated with the SRT in the TWOTALK condition [ $r = 0.75$ ,  $p = 0.013$ ]. Here, the correlation was stronger for the corresponding masking release [ $r = -0.77$ ,  $p = 0.009$ ], as depicted in Fig. 2.8(e). When controlling for the individual hearing thresholds at 1.5 kHz, the correlation with the SRT was marginal, while the correlation with the masking release remained significant [SRT:  $r = 0.61$ ,  $p = 0.08$ ; masking release:  $r = -0.67$ ,  $p < 0.05$ ]. Generally, the observed correlations were only slightly affected when the effect of absolute hearing thresholds was partialled out. To some degree, this can be attributed to the homogeneity of the HI group in terms of their hearing thresholds.

The finding of a correlation between the SRTs for the dichotic LATSSN masker and binaural low-frequency TFS processing seems reasonable in view of the results reported by Schubert and Schultz (1962) and Levitt and Rabiner (1967). They found that the release from masking for dichotic speech in noise ( $N_0S_{0.5ms}$  or  $N_0S_\pi$ ) was primarily determined by interaural time or phase disparity at low frequencies. Besides, in the present study binaural masked detection and dichotic speech reception depended in the same way on binaural integration: While they could be accomplished monaurally, use of the binaural information would give rise to better performance. While the LATSSN condition assessed the ability to take advantage of an interaural timing mismatch between target speech and noise interferer, performance in the TWOTALK background depended particularly on the ability to separate the target talker and the two interfering talkers. Hence, the correlations found between the SRTs for TWOTALK background and the measures of TFS processing support the hypoth-

## 40 2. Relations between frequency selectivity, TFS processing, and speech reception

esis of Zeng et al. (2005) that TFS cues might be utilized in talker separation in order to improve performance in listening situations with competing talkers. In this respect, the correlation between speech reception in the TWOTALK background (in terms of SRT and masking release) and FM detection performance at 1.5 kHz, observed here, may indicate a potential contribution of the second formant region (cf. Peterson and Barney, 1952) to talker identification and separation.

TFS processing was not correlated with SRTs (or masking releases) in the fluctuating backgrounds, SAM and RAM, neither for full-spectrum nor filtered speech. Hence, in contrast to Lorenzi et al. (2006), no evidence was found for a relation between TFS processing and dip listening. This discrepancy might have been due to the fact that the HI listeners in Lorenzi et al. (2006) had “flat” moderate hearing losses ( $\sim 50$  dB HL), whereas the HI listeners in the present study had “normal” hearing thresholds up to 1 kHz. Furthermore, Lorenzi et al. tested TFS processing with processed speech stimuli, which exhibited more complex TFS patterns than the tone stimuli used in the present study (with the exception of the uncorrelated noise maskers in the  $N_uS_0$  masked detection task).

A correlation between frequency selectivity and speech reception, as previously reported in literature (e.g., Dreschler and Plomp, 1985; Horst, 1987), was not observed here. However, these studies often included estimates of frequency selectivity at frequencies above 1000 Hz, while, in the present study, frequency selectivity was estimated only at 750 Hz. This may explain the absence of a correlation in the case of the full-spectrum speech, but not for the low-pass filtered speech. Another possible explanation, which might also account for the results with filtered speech, is that several impairment factors contributed to the observed speech reception deficits in complementary ways. Indeed, when the low-frequency slopes of the estimated filters and the monaural FMDTs at 1.5 kHz (with added AM) were considered as joint predictors in a multiple regression analysis, their combined effect on the monaural SRTs in the SSN and SAM conditions was significant [combined effect of filter slope and FMDT for SSN:  $F(2, 7) = 9.6$ ,  $p = 0.01$ ; for SAM:  $F(2, 7) = 8.5$ ,  $p = 0.01$ ]. The combined effect was less significant when the ERB instead of the filter slope was considered [for SSN:  $p = 0.04$ ; for SAM:  $p = 0.05$ ]. However, regression results that rely upon

such conjunctions of variables, rather than on strong primary correlations, should be viewed with caution, particularly in view of the small number of subjects.

### *Possible relations to aging*

One concern is that the NH listeners in the present study were, on average, younger than the HI listeners (median age 28 and 63 years, respectively). This raises the question of possible age effects, as previous studies have suggested a relation between aging and deficits in TFS processing as well as speech reception (e.g., Pichora-Fuller and Schneider, 1992; Strouse et al., 1998; Schneider et al., 2002; Ross et al., 2007). Indeed, subject HI<sub>1</sub>, who was the youngest of the HI listeners, performed better than the other HI listeners on the three tests of TFS processing, particularly lateralization and FM detection. However, apart from her age, HI<sub>1</sub> also differed in terms of etiology, as her hearing loss was due to hypoxia at birth. For the remaining HI listeners (53–74 years), dichotic masked detection was significantly correlated with age, while results for the other TFS tests were not [dichotic masked detection:  $r = 0.81$ ,  $p = 0.01$ ; lateralization:  $r = 0.36$ ,  $p = 0.37$ ; FM detection:  $r = 0.13$ ,  $p = 0.75$ ]. Hence, it cannot be excluded that part of the TFS deficits observed for the HI listeners could be related to aging. Ross et al. (2007) recorded cortical auditory-evoked responses to tones with dynamic changes in IPD. They found that the highest carrier frequency, at which responses to changes in IPD could be detected, declined with age. This indicates that aging might induce or potentiate a degradation in the processing of TFS at a peripheral or central auditory level, which is not reflected in the pure-tone hearing thresholds.

#### **2.8.2 Listeners with obscure dysfunction**

The two OD listeners showed deficits in frequency selectivity and binaural masked detection, which were comparable to those of the HI listeners. In the lateralization task they performed even more poorly than most of the HI listeners, showing substantial deficits, particularly with lateralization in background noise. However, in contrast to the HI listeners, who showed similar deficits on binaural lateralization and monaural FM detection, the OD listeners did not show as clear deficits in the FM detection task

## 42 2. Relations between frequency selectivity, TFS processing, and speech reception

as in the lateralization task. Since FM detection was assessed monaurally, one might conjecture that it was the non-tested ear that was actually limiting the lateralization performance. However, this can be excluded, as both ears were screened initially on FM detection and the worse ear was chosen for further testing. A possible reason for the poor binaural TFS performance of the OD listeners could be the large bandwidth differences between their ears. Colburn and Häusler (1980) suggested that the output of differing filters, given a diotic wideband input signal, would be partly uncorrelated at the two ears, resulting in lateralization blur. However, this explanation does not account for the observed poor performance in quiet and in dichotic (uncorrelated) noise. Hence, it seems that the TFS processing was affected at the stage of binaural integration rather than at a preceding monaural stage. Alternatively, even if the binaural TFS information was accurately integrated, it might not have been accessible at following stages of auditory processing.

The OD listeners showed rather small deficits in the reception of full-spectrum speech, but clear deficits in the reception of lowpass-filtered speech. These deficits might, at least partly, be attributable to the deficits in frequency selectivity and binaural TFS processing which were observed to a similar extent for both OD listeners. However, additional personality-related factors, such as an individual's underestimation of their own hearing ability (lack of “auditory confidence”) may be involved in the phenomenon of obscure (auditory) dysfunction. Considering the heterogeneity of the clinical group of OD patients (e.g., Saunders and Haggard, 1989; Zhao and Stephens, 2000) and the fact that the diagnosis of OD is solely based on a self-rated disability, the necessity for such factors is almost self-evident.

## 2.9 Possible underlying impairment mechanisms

Figure 2.7 illustrates that TFS processing was related neither to audibility nor to frequency selectivity, although deficits were found in all of the tests. One may speculate about possible impairment sites and mechanisms underlying these deficits. Damage to or loss of OHCs has been shown to result in a loss of sensitivity and frequency selectivity (e.g., Evans and Harrison, 1976; Liberman and Dodds, 1984), while damage to or loss of inner hair cells (IHCs) does not seem to have any substantial effect on

## 2.9 Possible underlying impairment mechanisms

43

sensitivity or tuning of the remaining intact IHCs (e.g., Wang et al., 1997). Hence, OHC damage might have been responsible for the deficits in frequency selectivity and their relation to absolute threshold observed here (cf. Moore et al., 1999).

Several factors might have contributed to the deficits in TFS processing. A loss of OHCs could have resulted in a reduced precision of phase locking (Woolf et al., 1981). However, this is controversial, as other studies did not find evidence for such phase-locking anomalies (e.g., Miller et al., 1997). Apart from this, Woolf et al. found the reduced phase locking to be related to elevated absolute thresholds, which was not observed for the TFS deficits in the present study. Also, a loss of OHCs might have altered the spatiotemporal response pattern of the BM. As suggested by Moore (1996) and Moore and Skrodzka (2002), this could have affected TFS processing if TFS information was extracted by cross-correlation of the outputs of different places along the BM (e.g., Deng and Geisler, 1987; Shamma, 2001; Carney et al., 2002). Since the present study assessed OHC integrity in terms of frequency selectivity only at a single frequency, this option cannot be ruled out here.<sup>8</sup>

Alternatively, through partial section of the auditory nerve (AN), it has been shown that a loss of AN fibers of up to 90% does not necessarily result in elevated pure-tone thresholds (e.g., Schuknecht and Woellner, 1953). Hence, the observed TFS deficits in regions of normal hearing might be attributable to damage to or loss of AN fibers or the innervated IHCs. A related possibility concerns the (monaural) enhancement of phase-locking synchrony to low-frequency tones that has been observed in the cochlear nucleus (e.g., Joris et al., 1994) and which might be reduced in impaired hearing. The alternative possibility, however, that a specific binaural processing stage, such as interaural coincidence detection, was affected in the HI listeners seems implausible given the clear correlation between the monaural and binaural TFS deficits found in these listeners.

<sup>8</sup> In section 4.4 of chapter 4, for example, it will be illustrated that the difference between CRTs at two remote places on the BM is not necessarily correlated with auditory filter bandwidth at a single frequency.

## 44 2. Relations between frequency selectivity, TFS processing, and speech reception

### 2.10 Summary

In addition to deficits in speech reception, deficits in frequency selectivity and in phase-locking-based TFS processing were observed for HI listeners, despite testing in regions of normal hearing. The observed TFS deficits were not related to reduced frequency selectivity. Monaural and binaural TFS deficits, however, were found to be related, suggesting that the binaural deficits might have been attributable to a monaural impairment factor. Background noise did not have a larger effect on TFS processing for the HI listeners than for the NH listeners: Although the acuity of TFS processing was decreased for the HI listeners, it seemed to be as robust to noise interference as for the NH listeners. SRTs in a two-talker background and in lateralized noise, but not in amplitude-modulated noise, were correlated with TFS-processing performance, suggesting that TFS information might be utilized in talker separation and spatial segregation.

The OD listeners showed deficits in frequency selectivity and in binaural, but not monaural, TFS processing. Compared with the NH listeners, their SRTs were particularly elevated for lowpass-filtered speech.

These findings on auditory deficits as well as preserved auditory abilities may serve as constraints for future models of the impaired auditory system. Furthermore, they may help in defining an auditory profile for listeners with impaired hearing.

# 3

---

## Estimation of cochlear response times using lateralization of frequency-mismatched tones<sup>†</sup>

---

Behavioral and objective estimates of CRTs and traveling-wave (TW) velocity were compared for three NH listeners. Differences between frequency-specific CRTs were estimated via lateralization of pulsed tones that were interaurally mismatched in frequency, similar to a paradigm proposed by Zerin (1969). In addition, derived-band ABRs were obtained as a function of derived-band center frequency. The latencies extracted from these responses served as objective estimates of CRTs. Estimates of TW velocity were calculated from the obtained CRTs. The correspondence between behavioral and objective estimates of CRT and TW velocity was examined. For frequencies up to 1.5 kHz, the behavioral method yielded reproducible results which were consistent with the objective estimates. For higher frequencies, CRT differences could not be estimated with the behavioral method due to principle limitations of the lateralization paradigm. The method may be useful for studying the spatiotemporal cochlear response pattern in human listeners.

### 3.1 Introduction

The cochlea responds to sound with a displacement wave that propagates on the BM from base to apex (e.g., Ruggero, 1994; Robles and Ruggero, 2001). This “traveling wave” (TW) serves to separate the tonal components of a sound by distributing their responses as distinctive spatial and temporal vibration patterns along the BM.

---

<sup>†</sup> This chapter is based on Strelcyk and Dau (2009b).

### 46 3. Estimation of CRTs using lateralization of frequency-mismatched tones

The wave reaches maximum amplitude at a particular point before slowing down and decaying rapidly. The lower the frequency of a sound, the further its wave propagates down the cochlea. Hence, each point along the cochlea has a CF to which it is most responsive. This tonotopic map is an important organizational principle of the primary auditory pathway and is preserved all the way to the auditory cortex (Clarey et al., 1992).

At the level of the auditory nerve, the frequency of a tone is encoded both spatially, by its CF location, and temporally, by the periodicity of the responses in the nerve fibers that innervate the CF (see Ruggero, 1992). Several studies have suggested that the extraction of spatiotemporal information, i.e., the combination of phase-locked responses and systematic frequency-dependent delays along the cochlea (associated with the TW), may be important in the context of pitch perception (e.g., Loeb et al., 1983; Shamma and Klein, 2000), loudness perception (Carney, 1994), localization (e.g., Shamma et al., 1989; Joris et al., 2006), speech formant extraction (e.g., Shamma, 1985; Deng and Geisler, 1987), and tone-in-noise detection (e.g., Carney et al., 2002). It has been proposed that a distorted spatiotemporal response might be, at least partly, responsible for the problems of HI listeners to process TFS information (e.g., Moore, 1996; Moore and Skrodzka, 2002; Buss et al., 2004). However, so far, empirical evidence for spatiotemporal information processing in humans is lacking since BM response patterns are difficult to monitor.

This study focused on one important component of the spatiotemporal BM response pattern: the cochlear response time (e.g., Don et al., 1993), which reflects the propagation delay of the TW. Consistent estimates of frequency-specific CRTs in human have been obtained using different objective noninvasive methods, such as measurements of compound action potentials (e.g., Eggermont, 1976), stimulus-evoked OAEs (e.g., Norton and Neely, 1987; Tognola et al., 1997), tone-burst-evoked ABRs (e.g., Gorga et al., 1988) and derived-band click-evoked ABRs (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978b; Eggermont and Don, 1980; Donaldson and Ruth, 1993; Don et al., 1993).

Early psychoacoustic attempts to estimate CRTs or TW velocity were motivated by von Békésy's (1933) observation that the perceived position of clicks, presented to both ears, varied systematically when low-frequency masking tones were presented



to one ear. Elaborating on this, Schubert and Elpern (1959) presented clicks in the presence of high-pass filtered noise with cutoff frequencies differing by half an octave between the two ears. The ITD that centered the unified percept at the midline was taken as an estimate of the difference in CRTs between the BM places corresponding to the noise cutoff frequencies in the two ears. However, the TW velocity derived from these CRT disparities was substantially larger than the TW velocity estimates obtained by means of the above mentioned objective methods (e.g., Donaldson and Ruth, 1993). As mentioned by Deatherage and Hirsh (1959) and Zerlin (1969), interaural loudness differences of the clicks might have influenced lateralization in the paradigms used by von Békésy and by Schubert and Elpern.

Instead of using click stimuli, Zerlin (1969) used pulsed tones that were interaurally mismatched in frequency. The ITD for which a centered percept was obtained was taken as an estimate of the difference in CRTs between the BM places corresponding to the different tone frequencies in the two ears. The derived TW velocities were in good agreement with objective estimates of TW velocity (cf. Donaldson and Ruth, 1993). However, as noted by Neely et al. (1988), the reliability of Zerlin’s estimates may be limited considering the difficulty of the psychoacoustic task and the fact that no further reports have been published since the original study in 1969.

If the lateralization of the interaurally mismatched tones reflected differences in CRTs, the paradigm would present a direct link between early cochlear disparities and perception. Hence, particularly in view of the high temporal acuity of binaural auditory processing which resolves ITD changes of less than  $10 \mu\text{s}$  (Yost, 1974), this behavioral paradigm might serve as a complement to the objective measures of CRT mentioned above. Furthermore, Zerlin’s paradigm bears a close relation to the concept of (across-ear) spatiotemporal processing. In both concepts, lateralization is supposed to be based on the comparison of information from mismatched frequency channels in the two ears. However, it is not clear if the lateralization in Zerlin’s paradigm is based on interaural level differences (in the envelope at onset/offset), interaural time differences (in the fine structure) or a combination of both. Buus et al. (1984) suggested that TFS information during the first tone cycles might play a role in the lateralization of mismatched tones at low frequencies. This was supported by Magezi

and Krumbholz (2008), who provided evidence that the binaural system can extract TFS information from interaurally mismatched frequency channels.

In the present study, behavioral estimates of CRT differences and TW velocity were obtained for three NH listeners, using a similar paradigm as the one used by Zerlin (1969). In order to minimize measurement variability due to subjective listener criteria, an adaptive procedure was used to determine the ITD that centered the unified percept. The influences of loudness balancing, tone presentation level and potential between-ear asymmetries on the CRT and TW-velocity estimates were examined. For direct comparison, estimates of CRTs and TW velocities for the same listeners were obtained from derived-band ABRs. Since these estimates provide an objective “reference”, they are presented first.

## 3.2 Auditory brainstem responses

### 3.2.1 Method

#### *Listeners*

The three female listeners were aged between 23 and 24 years and had audiometric thresholds better than 20 dB HL (ISO 389-8, 2004) at all octave frequencies from 125 to 8000 Hz and 750 to 6000 Hz.

#### *Stimuli*

Rarefaction clicks were produced by applying 83- $\mu$ s rectangular pulses (generated in MATLAB®) to an Etymotic Research ER-2 insert earphone. The clicks were presented monaurally at a level of 93 dB peak-to-peak equivalent sound pressure level (ppe SPL), with a repetition rate of 45 Hz. The acoustic clicks were calibrated using a B&K artificial ear (4157) with a B&K coupler (DP0370), a B&K microphone (4134), a B&K measuring amplifier (2636) and a Hewlett Packard digitizing oscilloscope (54503A). Response latencies were corrected for a constant 1-ms delay introduced by the tubing of the ER-2 earphone.

Ipsilateral pink-noise masking was used to obtain derived-band ABRs (Don and

Eggermont, 1978). High-pass noise maskers with cutoff frequencies of 0.5, 1, 2, 4 and 8 kHz were generated in the spectral domain as random-phase noise (with components outside the passband set to zero) and played back via a second ER-2 insert earphone, which was coupled to the first ER-2 earphone via an insert probe. The spectrum level of the high-pass noise maskers was identical to that of the broadband pink noise, for which a level of 91 dB SPL was found to be sufficient to mask the ABR to the 93-dB ppe SPL clicks.

Perceptual click thresholds were measured for 500-ms click trains using a 3I-3AFC task, tracking the 71%-correct point (one up, two down) on the psychometric function. The final threshold was estimated as the arithmetic mean over three runs.

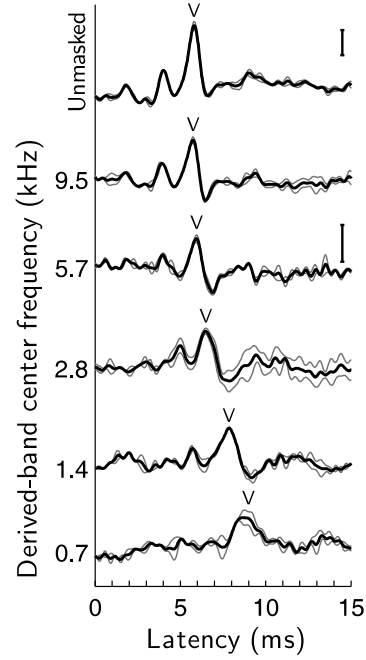
### ***ABR recordings***

Listeners were lying on a couch in an acoustically and electrically shielded booth. The ABRs were measured differentially between electrodes applied to the vertex ( $C_z$  in the 10/20 system) and the ipsilateral mastoid ( $M_1$  or  $M_2$ ). Another electrode applied to the forehead ( $F_{pz}$ ) served as ground. The electrode signals were acquired using a Neuroscan SynAmps 2 system, at a sampling rate of 20 kHz, and off-line bandpass filtered between 0.1 and 2 kHz (forward-backward filtering). Weighted averaging, as discussed in Elberling and Wahlgreen (1985) and in Don and Elberling (1994), was used for estimation of the auditory evoked potentials. Two replications, each consisting of 4096 sweeps, were recorded. The 4096 sweeps were subdivided into 16 equally sized blocks and averaged. Each block was weighted inversely proportional to its amount of background noise, which was estimated as the sweep-to-sweep variance at a single point in time (Elberling and Don, 1984). The residual background noise level in the final evoked potential estimates was 23 nV, averaged across listeners and conditions.

### ***Analysis***

Narrow-band cochlear contributions to the ABR were derived by means of the derived-band technique (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978a,b). Derived-band ABRs, i.e., differences between the ABR responses to clicks in adjacent

Figure 3.1: Examples of unmasked and derived-band ABR responses to 93-dB ppe SPL clicks from one listener. Two replications (gray) and their average (black) are shown. Wave V's are indicated by the corresponding symbols. The bars to the right represent 200 nV. If no bar is shown, the nearest bar above holds.



high-pass maskers, were obtained and the corresponding wave-V latencies were extracted. The center frequencies of the derived bands were computed as the geometric means of the two corresponding high-pass cut-off frequencies (Parker and Thornton, 1978b). The frequency of 11.3 kHz, where the click acoustic power was attenuated by 30 dB, was chosen as the upper frequency limit of the highest derived-band. Hence, the following frequencies were assigned to the derived bands: 0.7, 1.4, 2.8, 5.7 and 9.5 kHz.

Figure 3.1 illustrates a series of derived-band ABR responses from one listener. Wave V's are indicated. As can be seen, wave-V latencies increased with decreasing derived-band center frequency. For the further analysis of the wave-V latencies, the following latency model was adapted from Neely et al. (1988):

$$\tau(f) = a + bf^{-d}, \quad (3.1)$$

where  $f$  represents the derived-band center frequency, normalized to 1 kHz, and  $a, b$

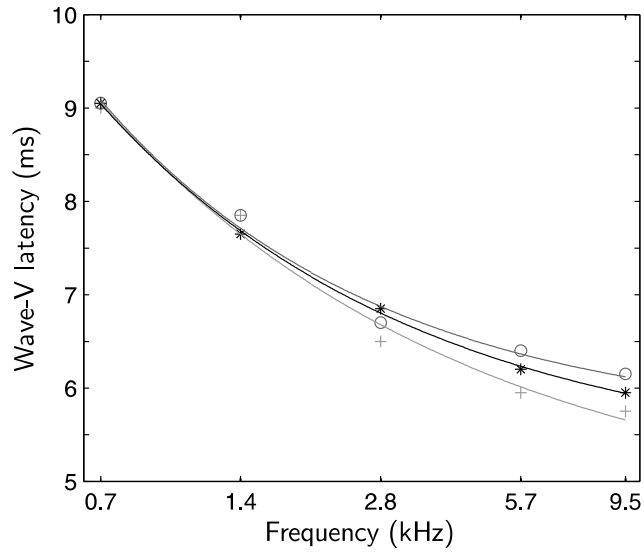


Figure 3.2: Measured derived-band ABR wave-V latencies (symbols) for three listeners in response to 93-dB ppe SPL clicks, as a function of the derived-band center frequency. The solid curves show individual model fits according to Eq. (3.1).

and  $d$  are fitting constants. The model parameter  $a$  represents an asymptotic delay. It reflects the post-cochlear contributions, i.e., synapse and neural conduction delays, to the wave-V latency, which are independent of frequency (cf. Don and Eggermont, 1978; Ponton et al., 1992; Ruggero, 1992).

### 3.2.2 Results

Figure 3.2 shows the measured (symbols) and fitted (curves) wave-V latencies. The results of all three listeners were similar. Latencies decreased with increasing frequency, in agreement with previous reports in literature (e.g., Eggermont and Don, 1980). The latency model specified in Eq. (3.1) provided a good description of the individual latency data, with a residual rms fitting error of 0.1 (SD 0.05) ms, averaged across listeners. The mean estimated parameters were:  $a = 5.1$  ms,  $b = 3.2$  ms and  $d = 0.6$ . The average perceptual click threshold was 33.7 (SD 2.5) dB ppe SPL.

### 3.3 Lateralization of mismatched tones

#### 3.3.1 Method

##### *Listeners*

The lateralization measurements were performed by the same listeners who participated in the ABR measurements.

##### *Stimuli and procedures*

Short trains of tone bursts with interaurally mismatched frequencies  $f_1$  and  $f_2$  were presented to the two ears, as illustrated in Fig. 3.3. In the following, the notation  $f_1|f_2$  is used where  $f_1$  represents the frequency of the tone presented in the ABR test-ear and  $f_2$  the frequency of the tone presented in the other ear. The considered tone frequencies were: 400|480 Hz, 800|900 Hz, 1000|900 Hz and 1400|1550 Hz. Each tone burst had a total duration of 40 ms, including an exponential onset with a rise time of 10 ms and a 10-ms raised-cosine shaped offset ramp. In contrast to Scharf et al. (1976) and Buus et al. (1984), who used exponential ramps at onset and offset, here a cosine offset-ramp was used in order to minimize spectral splatter. The tones were presented in sine phase, i.e., the onset ramp started with the positive-going zero crossing of the sinusoid. Each train consisted of six tone bursts, separated by 40-ms silent gaps. Its lateralization was varied by introducing a waveform delay to one of the ears, giving rise to an ITD. The ITD that produced a unified percept centered at the midline was measured.

A 2I-2AFC task was used. The first interval always contained the diotic reference tone-burst train, consisting of both tones (with frequencies  $f_1$  and  $f_2$ ) in both ears, while the second interval contained the  $f_1|f_2$  target train. Listeners were instructed to indicate if the latter was lateralized to the left or right side relative to the reference train. In order to ease the task, the whole trial consisting of reference and target train was repeated once before the listener made a response. If the target train was lateralized to the right, the ITD was adjusted such that the percept would move further to the left in the next presentation, and vice versa. Following the adaptive procedure for subjective judgments introduced by Jesteadt (1980), two sequences of trials were

### 3.3 Lateralization of mismatched tones

53

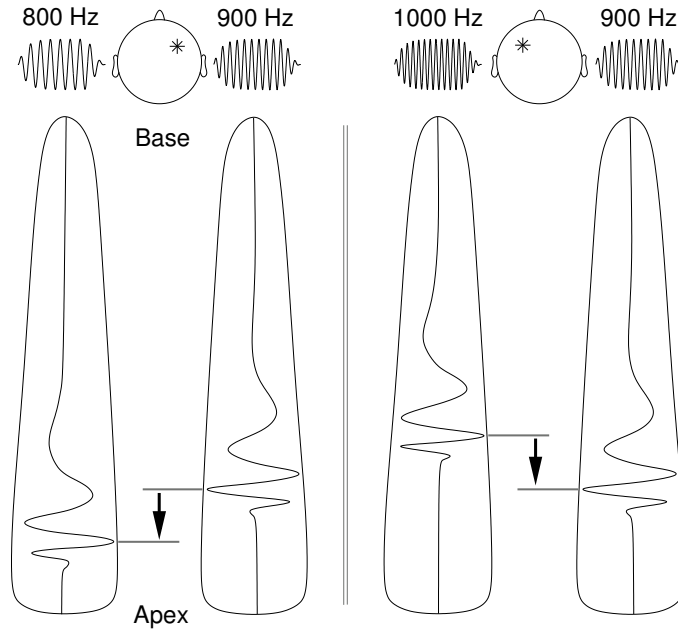


Figure 3.3: The stimuli used in the lateralization task, for the 800|900-Hz (top left) and 1000|900-Hz (top right) conditions (not in proportion). In the depicted configuration, the left ear is the ABR test-ear. BM traveling waves are indicated at the bottom. It is assumed that the CRT disparities, indicated by the arrows, can be measured in terms of the ITDs that center the percepts at the midline.

interleaved, tracking 71% (one up, two down) and 29% (two up, one down) lateralization to the right, respectively. Each of these sequences was terminated after ten reversals, and the tracked ITDs were estimated as the arithmetic means of all ITD values following the sixth reversals. Subsequently, the ITD yielding a centered percept was estimated by calculating the mean of the two ITDs leading to 71% and 29% lateralization judgments to the right.

ITDs were measured for tone levels of 50 and 75 dB SPL. In addition to the ITDs in quiet, for the 800|900-Hz tones at 75 dB, ITDs were measured in the presence of a diotic notched-noise background (flat-spectrum noise bands of 100–700 Hz and 1000–9000 Hz), which masked excitation spread to remote frequencies. The noise was

### 54 3. Estimation of CRTs using lateralization of frequency-mismatched tones

presented continuously during the whole run, with a spectrum level of 16 dB SPL. For higher levels, a fused position could no longer be perceived.

Prior to actual data collection, listeners received up to ten runs of training until consistent ITD results were obtained. The final ITD was estimated as the arithmetic mean over four interleaved runs. If the SD over these runs, relative to the mean ITD, exceeded a factor of 0.1, additional runs were taken and the average of all was used. The final relative standard error of the ITD estimate, averaged across listeners and conditions, was 0.05.

#### ***Loudness balancing***

In addition to the conditions where the tones were presented at equal SPLs, ITDs were measured with the tones balanced in loudness between the two ears. Loudness balancing was also applied by Zerlin (1969). The adaptive procedure introduced by Jesteadt (1980) was used for the loudness balancing of the frequency-mismatched tones. The first interval contained the  $f_1$ -tone, presented to the ABR test-ear, and the second interval contained the  $f_2$ -tone, presented to the other ear. Listeners were instructed to indicate if the second tone was perceived as softer or louder than the first tone. As in the lateralization task, the whole trial was repeated once before the listener made a response. The interaural level balance was adjusted to yield both 71% and 29% judgments of the second tone to be the louder one. The point of equal loudness was estimated as the arithmetic mean of these two loudness adjustments. An equal number of runs was performed with the opposite order of presentation, i.e., with the  $f_2$ -tone presented in the first interval and the  $f_1$ -tone presented in the second interval.

The final level adjustment for loudness balancing was estimated as the arithmetic mean over at least six interleaved runs. The final standard error of the level adjustment was 0.4 dB, averaged across listeners and conditions.

#### ***Apparatus***

The stimuli were generated in MATLAB® and converted to analog signals using a 24-bit digital-to-analog converter (RME DIGI96/8) with a sampling rate of 96 kHz. The stimuli were presented in a double-walled sound-attenuating booth via Sennheiser



HD580 headphones. Calibrations were done using a B&K artificial ear (4153) and, prior to playing, 128-tap linear-phase FIR equalization filters were applied to the stimuli, rendering the headphone frequency response flat.

### 3.3.2 Results and discussion

#### *Response-time differences*

The results of the lateralization measurements for the three listeners are presented in Table 3.1. It shows the ITDs that led to centered percepts of the 50-dB and 75-dB tones with interaurally mismatched frequencies  $f_1$  and  $f_2$ . The ITDs are stated for the conditions with and without interaural loudness balancing. As illustrated in Fig. 3.3, the frequency-mismatched tones with zero ITD were always lateralized towards the ear receiving the higher-frequency tone, consistent with previous reports in literature (e.g., von Békésy, 1963b; Zerlin, 1969). Hence, the sound presented to this ear required a delay in order to center the percept. The centering ITDs were generally consistent and well reproducible. Therefore, the standard errors of the ITD estimates were relatively small. For comparison, the objective ABR wave-V latency differences  $\Delta\tau_{\text{ABR}}$  are also represented in Table 3.1 (rightmost column). They were calculated on the basis of the individual latency fits to the derived-band ABR data, which followed the model in Eq. (3.1) and were shown in Fig. 3.2. Since these latency differences were very similar for the three listeners, only average  $\Delta\tau_{\text{ABR}}$  values are represented in the table. The lowest derived-band frequency was 700 Hz. Therefore, the extrapolation to lower frequencies (400|480 Hz) should be regarded with caution. At the remaining frequencies of 800|900 Hz, 1000|900 Hz and 1400|1550 Hz (second, third, and fourth rows in Table 3.1, respectively), the perceptual ITD-based measure and the objective ABR-based measure yielded very similar results. The average rms deviation between the ITDs (without loudness balancing) and the latency differences  $\Delta\tau_{\text{ABR}}$  was 39  $\mu\text{s}$ .

The ITDs reflect interaural time differences whereas the ABR latency differences  $\Delta\tau_{\text{ABR}}$  reflect monaural time differences. Hence, part of the remaining deviations between these two could be due to differences in CRTs between the left and right cochleae. Therefore, the ITDs for the 800|900-Hz and 1000|900-Hz tone pairs were added (see fifth row in Table 1). Since these tone pairs shared the common reference

Table 3.1: The ITDs yielding centered percepts of the tones with interaurally mismatched frequencies  $f_1$  and  $f_2$ , for three listeners (the numbers in parentheses represent standard errors). “LB” denotes loudness balancing. The ABR wave-V latency differences  $\Delta\tau_{ABR}$  between the frequencies  $f_1$  and  $f_2$  are also given (in parentheses; SD across listeners). The value in square brackets is based on an extrapolation beyond the range of measured frequencies. Conditions for which the listener could not perform the lateralization task are indicated by “NM” (not measurable). Dashes indicate combinations that were not measured.

Tone level	$f_1   f_2$ (kHz)	ITD ( $\mu s$ ) for $NH_1$		ITD ( $\mu s$ ) for $NH_2$		ITD ( $\mu s$ ) for $NH_3$		$\Delta\tau_{ABR}$ ( $\mu s$ )
		with LB	w/o LB	with LB	w/o LB	with LB	w/o LB	
50 dB	0.4 0.48	442(38)	340(40)	404(6)	396(8)	357(8)	335(5)	[609(36)]
	0.8 0.9	205(12)	264(13)	186(6)	232(10)	184(7)	207(3)	259(4)
	1.0 0.9	187(18)	185(13)	224(2)	262(3)	173(3)	180(9)	216(4)
	1.4 1.55	110(5)	138(9)	NM	129(22)	NM	NM	158(8)
	0.8   1.0	392(21)	449(18)	410(7)	494(11)	356(8)	387(9)	474(7)
75 dB	0.8 0.9	99(3)	—	79(5)	—	61(7)	—	—
	0.8 0.9	190(6)	—	183(2)	—	192(6)	—	—
	in noise	—	—	—	—	—	—	—

frequency of 900 Hz (cf. Fig. 3.3), the sum estimates the time difference between 800 and 1000 Hz in the ABR test-ear. Still, similar deviations from the ABR latencies as for the single-tone-pair ITDs were observed for these “monaural” time differences. Hence, the remaining deviations did not seem to be attributable to asymmetries between the left and right cochleae.

In addition to the measurements at 50 dB, for the 800|900-Hz tones, measurements were also performed at the higher tone level of 75 dB. For all listeners, ITDs were shorter at 75 dB than at 50 dB, by an average factor of 2.5. However, in the presence of the notched-noise masker, the ITDs obtained with the 75-dB tones were essentially identical to those obtained with 50-dB tones presented in quiet. This is consistent with the interpretation that tone excitation at the level of 75 dB spread to places with higher CFs than the nominal tone frequencies  $f_1$  and  $f_2$ . At these places, CRT differences were smaller as a consequence of the exponentially decreasing latency-frequency dependence (cf. Fig. 3.2). The notched noise masked excitation spread to remote frequencies and therefore gave rise to similar CRT differences as obtained at the lower tone level of 50 dB, where spread of excitation played a minor role.

Different stimuli, clicks versus tones, were used for the ABR recordings and the lateralization measurements, respectively. It seems reasonable to assume that stimulation at equal sensation levels results in similar levels of excitation, summed across the BM. The average sensation level of the 50-dB SPL tones was 12 dB lower than the average sensation level of the 93-dB ppe SPL clicks. However, both tones and clicks should have elicited roughly comparable amounts of excitation within the one-octave-wide derived-band regions on the BM. Also, remaining level differences should be of minor importance, since the 75-dB tones yielded very similar ITDs compared to the 50-dB tones when notched-noise masking was applied.

All three listeners had more difficulties with the lateralization task for the mid-frequency tones (1400|1550 Hz) than for the low-frequency tones. At 1400|1550 Hz, listener NH<sub>2</sub> could not consistently lateralize the mismatched tones when loudness balancing was applied, while listener NH<sub>3</sub> could not consistently lateralize the tones whether loudness balancing was applied or not. None of the listeners could perform the task reliably for frequencies above 1.5 kHz. Here, the sound image could not be

lateralized with reasonable precision. It was perceived as rather diffuse and often did not cross the midline.

### ***Loudness balancing***

For all tone pairs, ITDs changed consistently when loudness balancing was applied: The ITD increased (decreased) when the level of the higher-frequency tone was increased (decreased). The mean level adjustment was 0.7 (SD 0.4) dB, averaged across listeners and conditions. The ITDs obtained without loudness balancing seemed to match the objective latency differences  $\Delta\tau_{\text{ABR}}$  slightly better than the ones obtained with loudness balancing. The average rms deviations were 39  $\mu\text{s}$  and 66  $\mu\text{s}$ , respectively, excluding the 400|480-Hz data.

Assuming that the observed loudness imbalances were due to between-ear “gain” differences rather than within-ear variations in gain between the tone frequencies  $f_1$  and  $f_2$ , equal SPLs at the two ears would be more appropriate than equal loudness. A between-ear gain difference would affect the lateralization of the diotic reference stimulus and target stimulus in the same way. Hence, the ITD necessary for matching their positions would not be affected. As a consequence, since the reference stimulus was not balanced in loudness between the two ears, no loudness balancing should be applied to the target stimulus either.<sup>1</sup>

The ITDs obtained with and without loudness balancing were fairly comparable. This indicates that the perceived lateralization of the frequency-mismatched tones towards the ear with the higher-frequency tone was not simply a consequence of interaural gain differences, but most likely reflected differences in CRTs.

### ***Traveling-wave velocity***

Assuming that the centering ITDs and latency differences  $\Delta\tau_{\text{ABR}}$  reflected travel times on the BM, the corresponding TW velocities were estimated using the cochlear

<sup>1</sup> The reference stimulus was not balanced in loudness since, for matched-frequency tones, equal SPLs instead of equal loudness at the two ears would give rise to a percept centered at the midline. As discussed by Durlach et al. (1981), the binaural system adapts to between-ear gain differences in such a way that equal-SPL tones are perceived at the midline. In this way, the correlation of auditory perception with visual and tactile perception is maximized.

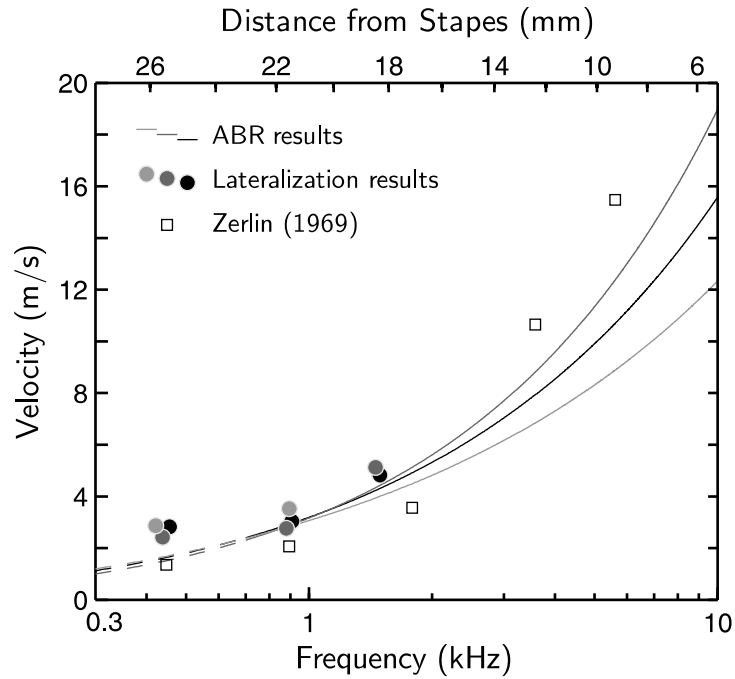


Figure 3.4: TW velocity as a function of frequency/distance from stapes for three listeners. The solid curves represent the individual velocity estimates derived from the derived-band ABR latencies. At low frequencies, the curves are dashed since they were extrapolated beyond the actual measurement range. The bullets denote the estimates based on the mismatched-tone ITDs. For better visibility, they are slightly horizontally displaced for the individual listeners. The squares are corresponding estimates based on the ITDs reported by Zerlin (1969).

frequency-place map supplied by Greenwood (1961).<sup>2</sup> Figure 3.4 shows the TW velocity estimates, based on the ABR latencies (curves) and the centering ITDs (bullets) obtained for the 50-dB tones without loudness balancing. The behavioral velocity estimate at 890 Hz (geometric mean of 800 and 1000 Hz) is based on the “monaural”

<sup>2</sup> The further assumption is made that the response time at a given CF place of the BM is the same for tonal stimulation with frequencies at and below this CF. This corresponds to constant group delays, i.e., constant slopes of the BM phase response (cf. Ruggero and Rich, 1987; Robles and Ruggero, 2001). For the mismatched 800|900-Hz tone pair, for example, the traveling waves in response to the 800-Hz and 900-Hz tones would reach the 900-Hz CF place at the same time. Hence, CRT differences would reflect the travel time between the 800-Hz and 900-Hz CF places.

### 60 3. Estimation of CRTs using lateralization of frequency-mismatched tones

time difference obtained by summing the 800|900-Hz and 1000|900-Hz ITDs as described above. For direct comparison, the open squares indicate velocities that were derived from Zerlin’s (1969) ITDs.<sup>3</sup>

The ITD-based velocity estimates were consistent with the ABR-based velocity estimates. In both measures, velocities increased with increasing frequency. In order to compare the ITD-based estimates at 440 Hz with the ABR-based estimates, the ABR data were extrapolated beyond the actual measurement range (dashed part of the curves). Here, the deviations between the two measures were larger than at the higher frequencies of 890 and 1470 Hz, reflecting the corresponding deviations of the CRT estimates (compare ITDs and  $\Delta\tau_{\text{ABR}}$  values in Table 3.1). The larger behavioral velocity estimates at 440 Hz might indicate that the actual latency-frequency functions were less steep at the low frequencies (below about 700 Hz) than the predictions based on the extrapolation of the ABR latencies (Fig. 3.2). This would be consistent with the latency-frequency curves in Neely et al. (1988, their Fig. 1), obtained from tone-burst-evoked ABRs, which showed shallower slopes for frequencies below about 500 Hz than for the higher frequencies.

Only small inter-individual differences were observed for frequencies up to 2 kHz, consistent with Donaldson and Ruth (1993). For frequencies above 1.5 kHz, no centering ITDs and thus no behavioral velocity estimates could be obtained in this study. At low frequencies, the velocity estimates were higher than the ones based on Zerlin’s ITDs (open squares).<sup>4</sup> Zerlin also estimated TW velocities at high frequencies. These velocities were larger than the velocities at low frequencies and roughly consistent with the present ABR-based estimates.

<sup>3</sup> Zerlin (1969) used the cochlear frequency-place map supplied by von Békésy (1963a) in order to derive TW velocities from the centering ITDs. In the present study, the Greenwood (1961) cochlear map was taken as a basis of all TW velocity estimates. Therefore, the TW velocities shown here were derived directly from the ITDs reported by Zerlin using the Greenwood map.

<sup>4</sup> These deviations cannot be attributed to the fact that Zerlin (1969) applied loudness balancing. Velocity estimates based on the ITDs obtained with loudness balancing (not shown) always fell in the same range or above the ones obtained without loudness balancing, but never below as Zerlin’s.

### 3.4 Overall discussion

The behavioral estimates of CRT and TW velocity, based on lateralization measurements, were consistent with the objective estimates based on ABR measurements. This is an interesting result, given the different experimental paradigms. It strongly supports the hypothesis that the ITDs that produced centered sound images reflected differences in CRTs between remote places on the BM. This hypothesis is corroborated by the observed influence of tone level: ITDs decreased with increasing level. This is consistent with an explanation in terms of spread of excitation on the BM and thus indicates that the perceived lateralization of the mismatched tones reflected cochlear disparities.

Despite the encouraging results of the lateralization paradigm for tone frequencies up to 1.5 kHz, no behavioral estimates of CRT could be obtained at higher frequencies. This was due to fundamental limitations in the lateralization paradigm, which are discussed in the following. In principle, a large frequency mismatch  $|f_2 - f_1|$  between the tones would be desirable to increase the accuracy of the ITD estimate. However, with increasing frequency mismatch, it becomes increasingly difficult to attribute a fused position (Scharf, 1972). More importantly, the lateralization threshold, i.e., the ITD for which the position of a non-centered sound object can just be distinguished from that of a centered object, increases strongly as soon as the interaural frequency mismatch exceeds a value that corresponds to the critical bandwidth for that frequency (Scharf et al., 1976; Buus et al., 1984). Scharf et al. found this bandwidth to be roughly independent of tone level and tone duration. The centering ITD, reflecting CRT disparity, needs to be larger than the corresponding lateralization threshold in order to be measurable. Therefore, in the present study, each tone pair was chosen such that the frequency mismatch between the tones did not exceed the critical bandwidth at the corresponding center frequency. The tone level of 50 dB SPL was comparable to the level of 50 phon used by Zerlin (1969). It was chosen as a compromise between decreasing lateralization thresholds and increasing spread of excitation with increasing tone level.

The feasibility of the measurements can, in principle, be predicted by comparing expected CRT disparities for maximally mismatched tones (tones that fall just within

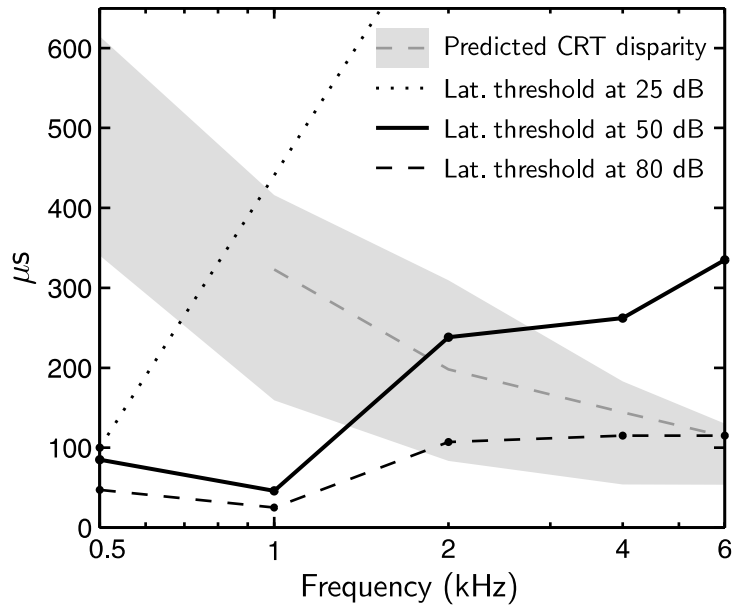


Figure 3.5: Predicted CRT disparities (gray) for maximally mismatched tones as a function of the center frequency of the tones. The gray shaded area indicates CRT disparities based on the TW velocity estimates given in Donaldson and Ruth (1993, section III.A.). The gray dashed curve shows disparity estimates based on the TW velocity estimates obtained in the present study (curves in Fig. 3.4). The black curves indicate the lateralization thresholds at 25 dB SPL (dotted curve), 50 dB SPL (solid curve), and 80 dB SPL (black dashed curve), obtained by Scharf et al. (1976) and Buus et al. (1984).

the same critical band) with the corresponding lateralization thresholds. First, critical bandwidths at 500, 1000, 2000, 4000, and 6000 Hz were extracted by digitizing the figures in Scharf et al. (1976). The obtained values were 115, 163, 310, 702, and 1080 Hz, respectively. In the next step, the frequencies of the maximally mismatched tones were calculated such that the geometric means of the two frequencies were equal to 500, 1000, 2000, 4000, and 6000 Hz. At 2000 Hz, for example, the tone frequencies were 1850 and 2160 Hz. Distances between the corresponding CF places on the BM were calculated based on the Greenwood (1961) frequency-place map. Then, expected minimal and maximal CRT disparities (“travel times”) were calculated for these BM distances, using the maximal and minimal TW-velocity esti-



mates, respectively, given in Donaldson and Ruth (1993, their Fig. 10). The predicted CRT disparities are shown in Fig. 3.5 (gray shaded area). The gray dashed curve indicates disparity estimates that are based on the average TW velocities obtained for the three listeners in the present study (curves in Fig. 3.4). As can be seen, CRT disparities for the maximally-mismatched tones decrease with increasing center frequency of the tones. Furthermore, the estimates based on TW velocities obtained in the present study are consistent with those based on the TW velocities in Donaldson and Ruth (1993). Figure 3.5 also shows the lateralization thresholds at the different tone levels of 25 dB SPL (dotted curve), 50 dB SPL (solid curve), and 80 dB SPL (black dashed curve), obtained by Scharf et al. (1976) and Buus et al. (1984).<sup>5</sup> Up to frequencies of about 1.5 kHz, the predicted CRT disparities are larger than the corresponding lateralization thresholds for mismatched 50-dB tones (solid curve) and are therefore measurable. However, with increasing frequency, the CRT disparities fall below the lateralization thresholds and are not measurable at a tone level of 50 dB. In theory, they are measurable using tone levels of about 80 dB and higher, since lateralization thresholds are smaller at these higher levels (black dashed curve). However, this assumes that the spread of excitation can be adequately confined, for example by means of notched-noise masking. For high frequencies of about 4 kHz or higher, the predicted CRT differences are too small to be measurable, even for tone levels of 80 dB. These predictions are consistent with the finding from the present study that ITDs could not be obtained for 50-dB tones at frequencies above 1.5 kHz.

The frequency mismatches for all tone pairs used by Zerlin (1969) exceeded the critical bandwidths given by Scharf et al. (1976) and Buus et al. (1984). For the 3200|4000-Hz and 5000|6300-Hz tone pairs, the reported centering ITDs clearly fall below the corresponding lateralization thresholds in those studies. Furthermore, Scharf et al. emphasized the importance of controlled tone-onset phases for tone frequencies below about 2 kHz: Without controlling the onset phase, their ITD data were inconsistent and the observed lateralization thresholds became substantially larger. Zerlin (1969), however, did not control onset phases. Hence, the validity of his results appears questionable both at low and high frequencies.

<sup>5</sup> The actual tone level at 500 Hz was 59 dB SPL, not 50 dB SPL.

64      3. Estimation of CRTs using lateralization of frequency-mismatched tones

---

One might argue that part of the discrepancies could be due to different ramp durations. Zerlin (1969) used 2.5-ms ramps, whereas 10-ms ramps were used in the present study as well as in Scharf et al. (1976) and Buus et al. (1984). However, even with such short ramp durations (tested in pilot measurements), it was not possible to obtain consistent ITD data at high frequencies. Apart from this, the percept gained a click-like character indicating a loss of frequency specificity.

In summary, due to the above limitations of the lateralization paradigm with frequency-mismatched tones, it is impossible to estimate CRT disparities (across remote BM places) at high frequencies with this method. However, for frequencies up to 1.5 kHz, the method yielded estimates of CRT disparities that were reasonably accurate, in terms of variability across measurements, and which were consistent with objective estimates. Hence, the lateralization method may be a valuable tool for studying aspects of the spatiotemporal cochlear response, particularly at low frequencies (below 500 Hz), where the accuracy of objective methods is limited.

## 4

---

### **Relation between derived-band auditory brainstem response latencies and frequency selectivity<sup>‡</sup>**

---

Derived-band click-evoked ABRs were obtained for NH listeners and sensorineurally HI listeners. The latencies extracted from these responses, as a function of derived-band center frequency and click level, served as objective estimates of CRTs. For the same listeners, auditory-filter bandwidths at 2 kHz were estimated using a behavioral notched-noise masking paradigm. Generally, shorter derived-band latencies were observed for the HI than for the NH listeners. Only at low click sensation levels, prolonged latencies were obtained for some of the HI listeners. The behavioral auditory-filter bandwidths accounted for the across-listener variability in the ABR latencies: CRT decreased with increasing filter bandwidth, consistent with linear-systems theory. The results link CRT and frequency selectivity in human listeners and offer a window to better understand how hearing impairment affects the spatiotemporal cochlear response pattern.

#### **4.1 Introduction**

Decoding of spatiotemporal information in the peripheral auditory system may be important for several auditory abilities, such as pitch perception (e.g., Loeb et al., 1983; Shamma and Klein, 2000), localization (e.g., Shamma et al., 1989), and speech formant extraction (e.g., Deng and Geisler, 1987). A distorted spatiotemporal response might be, at least partly, responsible for deficits in the processing of TFS information

---

<sup>‡</sup> This chapter is based on Strelcyk et al. (2009).

#### 66 4. Relation between derived-band ABR latencies and frequency selectivity

in HI listeners (e.g., Moore, 1996; Moore and Skrodzka, 2002; Buss et al., 2004). As discussed in chapter 2, this may be one of the reasons for their difficulties to understand speech in background noise. Hence, it is important to gain a better understanding of how hearing impairment affects the spatiotemporal behavior of the auditory periphery. In this study, CRT was investigated as it is an important component of the spatiotemporal response. Changes in CRT, due to cochlear hearing impairment, may result in distortions in the spatiotemporal response pattern.

The CRT can be considered as the sum of a cochlear transport time and a filter build-up time. It has been shown that concepts of linear-systems theory apply to some extent to BM responses (e.g., Goldstein et al., 1971; Geisler and Sinex, 1982; Recio and Rhode, 2000). In such a linear framework, the transport time corresponds to the signal-front delay in an auditory filter (Ruggero, 1980), while the filter build-up time corresponds to the duration from response onset to the time when the center of gravity (Goldstein et al., 1971; Ruggero, 1994) or peak amplitude (Geisler and Sinex, 1983) of the BM response is reached (see also Ruggero and Temchin, 2007). Don et al. (1998) suggested that the filter build-up time mainly reflects the delay which is introduced by the cochlear amplifier sharpening the BM tuning (e.g., Robles and Ruggero, 2001).

CRTs have been studied extensively in NH listeners, using noninvasive methods such as measurements of compound action potentials (e.g., Eggermont, 1976), stimulus-evoked OAEs (e.g., Norton and Neely, 1987; Tognola et al., 1997), tone-burst-evoked ABRs (e.g., Gorga et al., 1988) and derived-band click-evoked ABRs (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978b; Eggermont and Don, 1980; Don et al., 1993).

Apart from studies on Ménière’s disease (e.g., Eggermont, 1979; Rutten, 1986; Donaldson and Ruth, 1996), where changes in CRT are supposed to reflect changes in cochlear transport time due to endolymphatic hydrops (Thornton and Farrell, 1991), only few studies examined CRT in HI listeners. Donaldson and Ruth (1996) measured derived-band ABRs in HI listeners and found no alterations as a consequence of hearing loss (in the group without Ménière’s disease). In contrast, Don et al. (1998), using a similar method, reported a tendency towards shorter response latencies with increasing hearing loss. Since hearing loss is often related to reduced frequency selectivity in HI listeners (e.g., Tyler et al., 1983; Moore, 1996; Baker and Rosen, 2002),

Don et al. suggested that the shorter latencies reflected increased auditory-filter bandwidths, consistent with the uncertainty principle of Fourier analysis (Papoulis, 1962).

Animal studies provided empirical evidence for such a relation between CRT and frequency selectivity (e.g., Goldstein et al., 1971; Geisler and Sinex, 1983). However, in humans, this relation has not yet been demonstrated directly. Shera et al. (2002) estimated CRTs (in terms of BM group delays) from stimulus-frequency OAEs in humans. Based on relations between CRT and auditory-filter bandwidth from animal data, they predicted human filter bandwidths. The average bandwidths as a function of frequency were consistent with behavioral bandwidth estimates obtained in a forward-masking paradigm. However, the method by which Shera et al. (2002) obtained their CRT estimates is a matter of debate, as it relies on assumptions of how stimulus-frequency OAE delays relate to BM delays (Ruggero and Temchin, 2005). Furthermore, Ruggero and Temchin (2005) questioned the use of forward-masking filter bandwidths as measures of cochlear frequency tuning in the above study.

In the present study, CRTs were estimated from derived-band ABRs of NH and HI listeners, as a function of frequency and level. Possible alterations in CRT, linked with cochlear hearing impairment, were examined. In order to study explicitly the relation between CRT and frequency selectivity across the individual listeners, behavioral estimates of filter bandwidth were obtained at a frequency of 2 kHz, using a notched-noise simultaneous-masking paradigm (e.g., Patterson and Nimmo-Smith, 1980). Specifically, it was examined whether individual filter bandwidths could account for the observed across-listener variability in CRTs. It was expected that the across-listener variability within the group of HI listeners would provide valuable information when investigating the relation between CRT and frequency selectivity.

## 4.2 Auditory brainstem responses

### 4.2.1 Method

#### *Listeners*

The five NH listeners were aged between 23 and 25 years and had audiometric thresholds better than 20 dB HL (ISO 389-8, 2004) at all octave frequencies from

68 4. Relation between derived-band ABR latencies and frequency selectivity

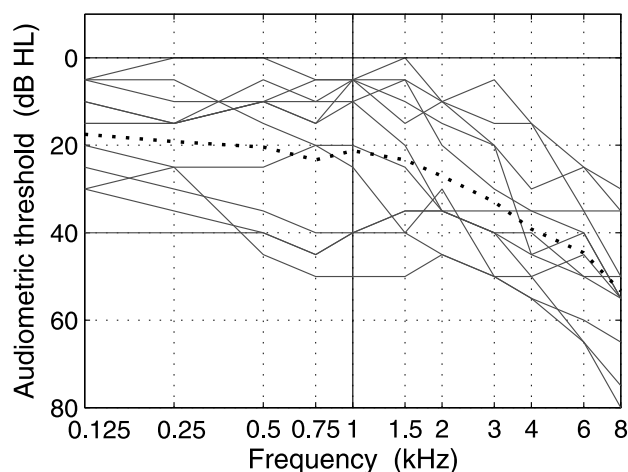


Figure 4.1: Audiograms of the twelve HI listeners. The dotted curve shows the arithmetic mean.

125 to 8000 Hz and 750 to 6000 Hz (three of them had already participated in the study presented in chapter 3). Only female listeners participated in the present study to avoid gender-related differences in the ABR latencies (e.g., Don et al., 1993). The twelve HI listeners were aged between 42 and 80 years (median: 55). Their better ears in terms of audiometric thresholds were chosen for further testing, so that effects of cross-ear listening could be ruled out. Their audiograms are shown in Fig. 4.1. All HI listeners showed sloping audiograms with a mild-to-moderate hearing loss at high frequencies. This is reflected in the mean audiogram, indicated by the dotted curve in Fig. 4.1. Audiograms were “normal” (thresholds  $\lesssim 20$  dB HL) up to 1 kHz in seven of the HI ears, while one audiogram reflected a mild hearing loss and four reflected a moderate hearing loss of up to 50 dB HL at these frequencies. The sensorineural origin of the hearing losses was established by means of otoscopy, bone-conduction and acoustic-reflex measurements. There were no indications of Ménière’s disease such as episodic vertigo, fluctuating hearing loss or the sensation of fullness. ABRs to 100-dB ppe SPL clicks were measured in both ears of all HI listeners. Since the interaural wave-V delay, the interaural wave-V amplitude difference, and the monau-

## 4.2 Auditory brainstem responses

69

ral wave I–V interpeak delays were within normal range for all ears, there was no indication of eighth-nerve tumors or brainstem lesions (Don and Kwong, 2002).

### *Stimuli*

Generally, the same stimuli and equipment as described in section 3.2.1 were also used in this study. Differences are stated in the following. The rarefaction clicks were presented monaurally at five equally spaced levels ranging from 16 dB SL (for one HI listener a lower level of 11 dB was used) to the upper fixed level of 93 dB ppe SPL. The number of intermediate levels was reduced if a small individual dynamic range would have resulted in a level spacing smaller than 5 dB. A broadband pink-noise level of 91 dB SPL was found to be sufficient to mask the ABR to the 93-dB ppe SPL clicks. For the lower click levels, the noise was attenuated with the click to maintain a fixed click-to-noise ratio.

The 93-dB ppe SPL clicks were presented in quiet (unmasked) and in five high-pass noise maskers, with cutoff frequencies of 0.5, 1, 2, 4 and 8 kHz. The frequency bands between 1 and 2 kHz as well as 2 and 4 kHz have previously been found to yield the most salient derived-band ABRs (e.g., Eggermont and Don, 1980). Therefore, in order to save measurement time, only the conditions with 1, 2 and 4-kHz high-pass maskers were measured for the lower click levels.

Perceptual click thresholds were measured for 500-ms click trains using a 3I-3AFC task, tracking the 71%-correct point (one up, two down) on the psychometric function. The final threshold was estimated as the arithmetic mean over three runs.

### *ABR recordings*

The same setup and procedures as described in section 3.2.1 were used.

### *Analysis*

Derived-band ABRs (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978a,b) were obtained and the corresponding wave-I and wave-V latencies were extracted. The following frequencies were assigned to the derived bands: 0.7, 1.4, 2.8, 5.7

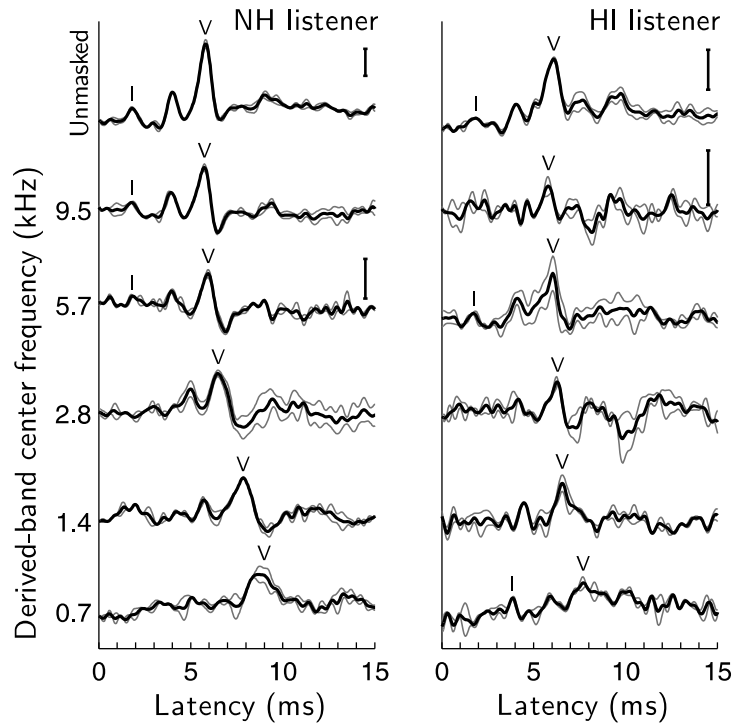


Figure 4.2: Examples of unmasked and derived-band ABR responses to 93-dB ppe SPL clicks from one NH listener (left) and one HI listener (right). Two replications (gray) and their average (black) are shown. Waves I and V are indicated by the corresponding symbols. The bars to the right represent 200 nV. If no bar is shown, the nearest bar above holds.

and 9.5 kHz. Figure 4.2 illustrates a series of derived-band ABR responses to 93-dB ppe SPL clicks from one NH listener (left) and one HI listener (right). Waves I and V are indicated. As can be seen, wave I's could not be identified for all derived bands. For both listeners, wave-V latencies increased with decreasing derived-band center frequency. However, the increase was larger for the NH listener than for the HI listener: the latter showed shorter latencies in the three lowest derived bands than the NH listener.

For the further analysis of the wave-V latencies, in terms of nonlinear statistical



modelling, the following latency model was adapted from Neely et al. (1988):

$$\tau(f, i) = a + b c^{0.93-i} f^{-d}, \quad (4.1)$$

where  $f$  represents the derived-band center frequency, normalized to 1 kHz,  $i$  represents click stimulus level, normalized to 100 dB ppe SPL, and  $a, b, c$ , and  $d$  are fitting constants. This parametrization slightly deviates from that in Neely et al. (1988). In contrast to their reference level of 0 dB, a reference level of 93 dB was chosen here, so that the parameter  $b$  reflects the forward latency at a frequency of 1 kHz and a click level of 93 dB. The model parameter  $a$  represents an asymptotic delay which is independent of frequency and level.

## 4.2.2 Results and discussion

### *Wave-V latency as a function of frequency*

Figure 4.3 shows the measured (symbols) and fitted (curves) wave-V latencies at 93 dB ppe SPL as a function of frequency for the NH (black) and the HI (gray) listeners. The latency model specified in Eq. (4.1) provided a good description of the individual latency data, with a residual rms fitting error of 0.17 (SD 0.09) ms, averaged across all listeners.

Latencies decreased with increasing frequency. The latencies for the NH listeners were consistent with previous reports in literature (e.g., Eggermont and Don, 1980; Don et al., 1993), both in absolute terms as well as frequency dependence. For eight of the twelve HI listeners, shorter latencies were observed than for the NH listeners (at 1.4 and 2.8 kHz). A similar trend was reported previously by Don et al. (1998). Also, the latency differences between wave V's of the 1.4- and 2.8-kHz derived bands were smaller for the HI than for the NH listeners [ $p = 0.05$ ]. The across-listener variability was larger among the HI listeners than the NH listeners, and the variability among the HI listeners was larger at the lower than at the higher frequencies, in terms of the wave-V latencies. At the lowest frequency of 700 Hz, a wave-V latency could not be identified for one of the HI listeners, whereas at the highest frequency of 9500 Hz, latencies could not be identified for four of the HI listeners.

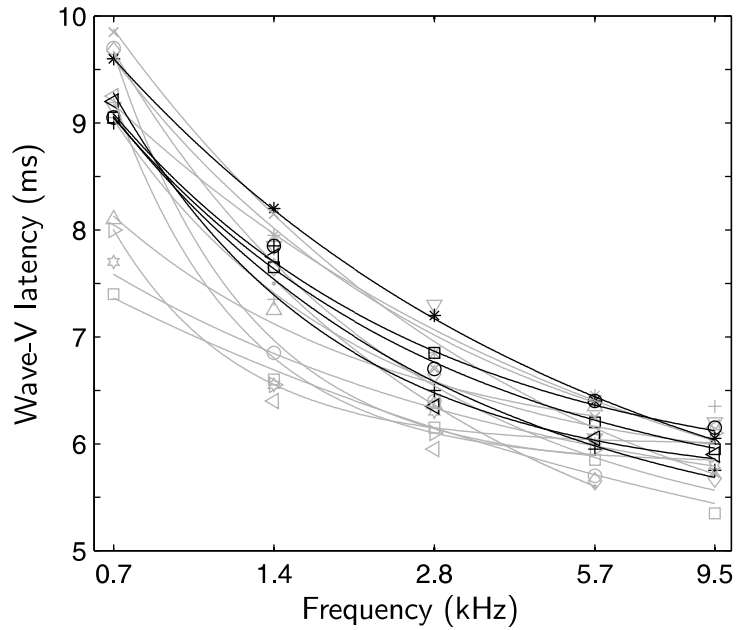


Figure 4.3: Measured derived-band wave-V latencies in response to the 93-dB ppe SPL clicks, as a function of derived-band center frequency, for the NH (black symbols) and HI listeners (gray symbols). The curves show individual fits, according to the latency model in Eq. (4.1).

### ***Wave-V latency as a function of click level***

Figure 4.4 shows the measured (symbols) and fitted (curves) wave-V latencies as a function of click level in the 2.8-kHz derived band. The results obtained for the 1.4-kHz derived band followed the same trends and are therefore not shown. Latencies decreased with increasing click level, consistent with the results from Eggermont and Don (1980) and Don et al. (1993). At high click levels (above about 80 dB ppe SPL), the HI listeners (gray curves) showed a trend to shorter latencies than the NH listeners (black curves), consistent with the trend in the data from Fig. 4.3, observed at 93 dB. At lower click levels, however, some of the HI listeners showed longer latencies than the NH listeners. This finding will be discussed further below.

To accommodate the repeated-measures design, the statistical analyses were car-

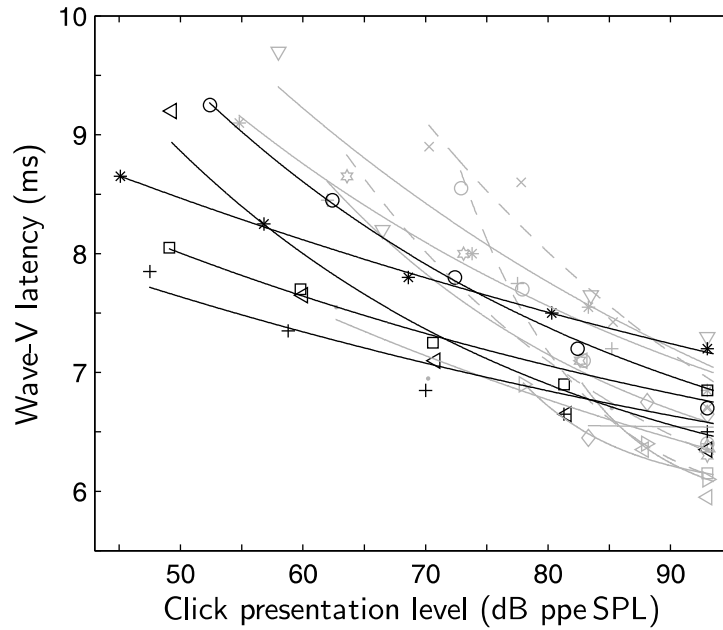


Figure 4.4: Measured derived-band wave-V latencies as a function of absolute click level for the 2.8-kHz derived band, for the NH (black symbols) and HI listeners (gray symbols). The curves show individual model fits.

ried out using nonlinear MEMs (Lindstrom and Bates, 1990; Pinheiro and Bates, 2000), as implemented in S-PLUS®. The between-listener variability that was not explained in terms of the fixed effect listener group was accounted for in terms of listener-specific random effects. An ANOVA was performed based on a nonlinear MEM which followed the latency model given in Eq. (4.1). The ANOVA confirmed the significance of derived-band frequency [ $F(1, 148) = 58.1, p < 0.0001$ ] and click level [ $F(1, 148) = 62.6, p < 0.0001$ ] on wave-V latency. Also the effect of listener group was significant [ $p < 0.001$ ], with a smaller parameter  $b$  and a steeper level slope  $c$  for the HI than for the NH listeners. Since the parameter  $b$  varied significantly across listeners [ $p < 0.0001$ ], it was modelled as listener-specific random effect. The estimated mean parameters were:  $a = 4.7$  ms,  $b = 3.4$  ms,  $c = 5.2$  and  $d = 0.50$ .

For some of the HI listeners, longer latencies were observed than for the NH lis-

#### 74 4. Relation between derived-band ABR latencies and frequency selectivity

teners (Fig. 4.4), for click presentation levels below about 80 dB ppe SPL. This cannot be explained in terms of cochlear filter bandwidth since the observed longer latencies would imply an implausible better-than-normal frequency tuning for the HI listeners. OHC damage might not only affect the frequency tuning and sensitivity of the BM (e.g., Evans and Harrison, 1976; Liberman and Dodds, 1984) but also its local stiffness. Donaldson and Ruth (1996) and Don et al. (1998) suggested that a loss of OHCs could result in decreased stiffness of the cochlear partition. Decreased stiffness would result in longer transport times and consequently in prolonged wave-V latencies. However, since the transport time itself reflects a “passive” BM property (Ruggero and Temchin, 2007), such changes in transport time would be expected to be independent of stimulus level. Hence, this explanation is unlikely to account for the fact that steeper slopes of the latency-level curves were observed for the HI than for the NH listeners.

The latency prolongations might be a result of reduced sensitivity reflected in the hearing losses of the HI listeners: They showed significantly higher click thresholds than the NH listeners [NH: 32.7 (SD 2.7) dB ppe SPL; HI: 54.3 (SD 13.5) dB ppe SPL;  $F(1, 15) = 12.3$ ,  $p = 0.003$ ]. Indeed, when latencies are plotted as a function of the click sensation level, as shown in Fig. 4.5, the latencies for the HI listeners fall well into or below the range of the NH listeners. Reduced BM sensitivity would result in decreased input levels to the IHCs. Recordings from AN fibers in response to tones have shown that AN first-spike latencies depend on stimulus level (see Heil, 2004). This has been attributed to temporal integration processes in the synapses between IHCs and AN fibers (Heil and Neubauer, 2003). Hence, increasing synaptic delays with decreasing input level might have been responsible for the abnormally long latencies observed in some of the HI listeners (when considered as a function of absolute SPLs).

Also, damage to or loss of IHCs might partly account for the latency prolongations. Synaptic delays, for example, would increase more strongly with decreasing stimulus level in less sensitive neurons, i.e., neurons with high spike thresholds (cf. Heil, 2004, Fig. 1). IHC loss does not necessarily result in elevated pure-tone thresholds (e.g., Schuknecht and Woellner, 1953) and is even less likely to be reflected

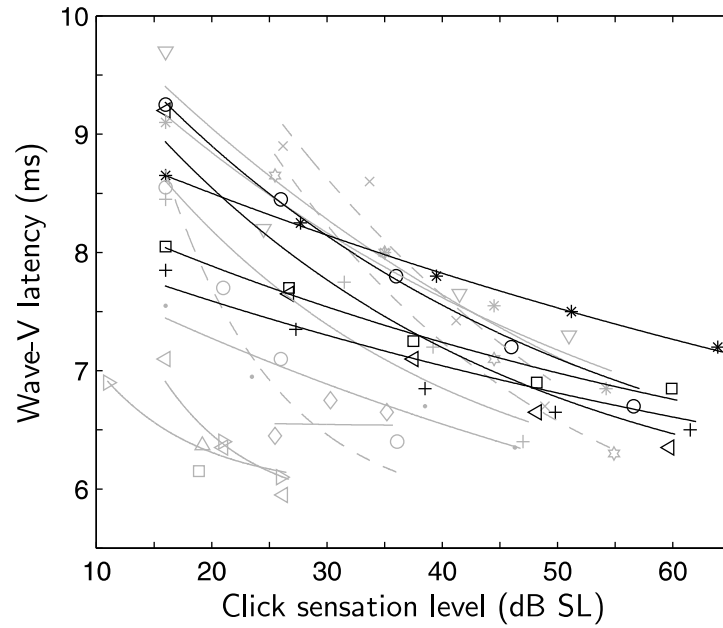


Figure 4.5: Same as Fig. 4.4, but as a function of click sensation level (click level re perceptual click threshold).

in thresholds for broadband clicks. Hence, the steep slopes of the latency-level curves for some of the HI listeners might be partly attributable to hidden IHC losses.

In order to test the tenability of this hypothesis, detection thresholds for sinusoidal 2-Hz frequency modulation (FMDTs) were measured for three of the HI listeners who showed steep latency-level curves (dashed curves in Fig. 4.4 and 4.5). Low-rate FM detection has been assumed to be based on TFS processing in terms of phase locking (e.g., Moore and Sek, 1996) and, as such, to be a measure of IHC or AN-fiber integrity (e.g., Buss et al., 2004). The FMDTs were measured using a carrier frequency of 1500 Hz, for which NH reference data were available from a previous experiment (chapter 2; see section 2.7.1 for a detailed method description). As illustrated in Fig. 4.6, the three tested HI listeners (gray bullets) showed markedly increased FMDTs compared with the six NH listeners (black squares) from chap-

76 4. Relation between derived-band ABR latencies and frequency selectivity

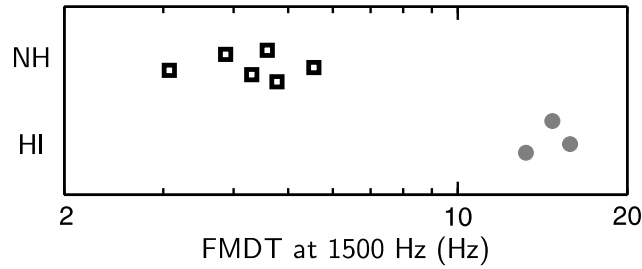


Figure 4.6: FMDTs at a carrier frequency of 1500 Hz, for six NH listeners (replot of data from Fig. 2.6, sixth column) and three of the HI listeners.

ter 2. This outcome is consistent with the hypothesis that steeply sloping latency-level curves might, at least partly, be linked to an IHC loss.

### *Estimation of cochlear response time*

In addition to CRTs, wave-V latencies reflect the delay introduced by the IHC-AN synapses and a central conduction time to the point in the brainstem which is responsible for the wave-V peak activity. Central conduction time has been shown to be independent of frequency (Don and Eggermont, 1978; Ponton et al., 1992; Don et al., 1993) and click level (Eggermont and Don, 1980). Synaptic time delays are independent of frequency (cf. Ruggero, 1992), but they may depend on stimulus level, as discussed above.

In order to estimate central conduction times, wave I–V interpeak delays were extracted from the derived-band responses (cf. Fig. 4.2). When interpeak delays were available for several derived bands and click levels, the average was taken as an estimate of the central conduction time in the individual listener. For the synaptic delay, a constant value of 0.8 ms was assumed (Eggermont, 1979). The difference between the wave-V latency and the sum  $\tau_{IV+0.8}$  of conduction time and synapse delay was taken as an estimate of the CRT (cf. Don et al., 1998). The average value of  $\tau_{IV+0.8}$  was 4.9 (SD 0.2) ms, with no significant difference between listener groups [ $p \sim 0.3$ ]. This value was also consistent with previous estimates of the asymptotic delay in literature (e.g., Neely et al., 1988; Donaldson and Ruth, 1993). The individual

## 4.2 Auditory brainstem responses

77

$\tau_{IV+0.8}$  delay was substituted for the asymptotic delay  $a$  in the model Eq. (4.1).<sup>1</sup> This yielded the new latency model

$$\tau(f, i) = \tau_{IV+0.8} + b c^{0.93-i} f^{-d}. \quad (4.2)$$

A new MEM, based on this model, confirmed the results of the previous MEM, in terms of parameter estimates as well as significance of derived-band frequency, click level and listener group. Also in terms of goodness of fit the two models were roughly equivalent, yielding equally good descriptions of the latency data.

Synaptic delays might not be constant, as assumed above, but increase with decreasing click level. In this case, the CRT estimates [second summand in Eq. (4.2)] would partly reflect the level-dependent synaptic delays and CRTs would be overestimated at low click levels. In the following, the relation between CRTs and behavioral estimates of frequency selectivity will be explored, in order to test if the across-listener variability in CRTs could be attributed to differences in frequency selectivity. While decreased frequency selectivity is expected to result in decreased CRTs, decreased sensitivity (and thus potentially decreased input levels at the synapse) should result in increased CRT estimates. Hence, the assumption of a constant synaptic delay is conservative with regard to the hypothesized relation between CRT and frequency selectivity.

<sup>1</sup> Alternatively, instead of using the  $\tau_{IV+0.8}$  delay, the asymptotic delay  $a$  in Eq. (4.1) could have been estimated directly from the individual wave-V latencies. However, this was problematic since wave-V latencies at 9.5 kHz (and 93 dB) were missing for four of the HI listeners. Even if these latencies had been available, it is questionable that the upper frequency limit of 9.5 kHz was sufficiently high to estimate the asymptotic delay. Subtraction of the wave I–V delay from the observed 9.5-kHz wave-V latency at 93 dB yielded on average a remaining delay of 1.9 (SD 0.1) ms for the NH and 1.7 (SD 0.4) ms for the HI listeners. It seems that this remainder cannot be solely accounted for by a synaptic delay, for which a value of about 0.8 ms is commonly assumed at comparably high stimulus levels (e.g., Eggermont, 1979; Robles and Ruggero, 2001). In parts, the remainder may reflect a finite CRT at the frequency of 9.5 kHz and the level of 93 dB.

## 4.3 Frequency selectivity

### 4.3.1 Method

#### *Listeners*

The notched-noise masking measurements were performed by the same listeners who participated in the ABR measurements.

#### *Stimuli and procedures*

Auditory filter shapes at 2 kHz were determined for the ABR test-ears using a notched-noise masking paradigm (cf. Patterson and Nimmo-Smith, 1980). The 2-kHz target tones of 440-ms duration were temporally centered in the 550-ms noise maskers. Maskers and tones were gated with 50-ms raised-cosine ramps. The noise was generated in the spectral domain as fixed-amplitude random-phase noise. Five symmetric ( $\delta f/f_0$ : 0.0, 0.1, 0.2, 0.3, and 0.4) and two asymmetric notch conditions ( $\delta f/f_0$ : 0.2|0.4 and 0.4|0.2) were used, where  $\delta f$  denotes the spacing between the inner noise edges and the signal frequency  $f_0$ . The outside edges of the noise maskers were fixed at  $\pm 0.8f_0$ . The tones were presented at a fixed level of 40 dB SPL for the NH listeners while the masker level was varied adaptively. For some of the HI listeners, a tone level of 40 dB would have resulted in a sensation level of less than 15 dB. In order to obtain reliable filter estimates, in these cases, the tone level was increased to ensure a minimum sensation level of 15 dB. The average tone level for the HI listeners was 47 (SD 8) dB.

A 3I-3AFC weighted up-down method (Kaernbach, 1991) was applied to track the 75%-correct point on the psychometric function. A run was terminated after 14 reversals. The threshold was defined as the arithmetic mean of all masker levels following the fourth reversal. Following a training run for each notch condition, the threshold was estimated as the average over three runs. If the SD of these three runs exceeded 1 dB, one or two additional runs were taken and the average of all was used.



### *Apparatus*

The stimuli were generated in MATLAB® and converted to analog signals using a 24-bit digital-to-analog converter (RME DIGI96/8), with a sampling rate of 48 kHz. The stimuli were presented in a double-walled sound-attenuating booth via Sennheiser HD580 headphones. Calibrations were done using a B&K artificial ear (4153) and, prior to playing, 128-tap linear-phase FIR equalization filters were applied, rendering the headphone frequency response flat.

### *Filter fitting*

A nonlinear minimization routine was implemented in MATLAB® to find the best-fitting roex filter in the least-squares sense, assuming that the signal was detected using the filter with the maximum SNR at its output. The  $\text{roex}(p, r)$  filter model (Patterson et al., 1982) and a more complex variant, the  $\text{roex}(p, w, t, p)$  model as used by Oxenham and Shera (2003), were considered. At the low-frequency side, the filter shape  $W(f)$  of the  $\text{roex}(p, w, t, p)$  filter is defined by

$$W(f) = (1 - w)(1 + pg) \exp(-pg) + w(1 + pg/t) \exp(-pg/t), \quad (4.3)$$

where  $g$  represents the deviation from the center frequency as a proportion of the center frequency,  $p$  determines the passband-slope of the filter,  $t$  determines the factor by which the tail-slope is shallower than the passband-slope and  $w$  determines the relative weights of the two slopes. The high-frequency side of the filter is described by a single slope,

$$W(f) = (1 + pg) \exp(-pg), \quad (4.4)$$

which is independent of the low-frequency side.

The ERB was computed as a measure of filter tuning.<sup>2</sup> Also, the uncertainty of

<sup>2</sup> Besides the ERB as a measure of filter tuning, also the 10-dB bandwidths were considered. However, because they yielded very similar results, for ease of comparison only the ERB results will be discussed further.

#### 80 4. Relation between derived-band ABR latencies and frequency selectivity

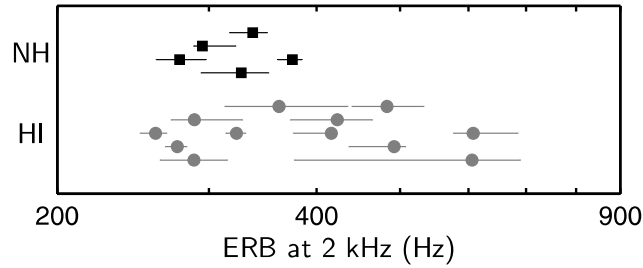


Figure 4.7: ERB of the  $\text{roex}(p, w, t, p)$  filter estimates at 2 kHz for the NH (black squares) and HI (gray bullets) listeners. The errorbars represent the 15th and 85th percentiles, which were estimated via bootstrapping; see text for details.

the ERB was estimated via bootstrapping: Based on the empirical standard errors of the individual notched-noise thresholds, for each listener, a large number of threshold curves was resampled. Subsequently, auditory filters were fitted to these threshold replications. The resulting bandwidth distribution yielded a confidence interval for the ERB.

#### 4.3.2 Results and discussion

The  $\text{roex}(p, w, t, p)$  filter model provided a good description of the individual notched-noise threshold data, with a residual rms fitting error of 0.5 (SD 0.3) dB, averaged across all listeners. The rms fitting error for the simpler  $\text{roex}(p, r)$  filter model was on average larger by a factor of 1.3 (1.0, 1.6); the numbers in brackets represent the 15th and 85th percentiles, respectively. Therefore, only the results for the  $\text{roex}(p, w, t, p)$  model will be discussed in the following. However, the pattern of results and conclusions would remain unchanged if the  $\text{roex}(p, r)$ -results were considered instead.

Figure 4.7 shows the estimated ERBs for the NH (black squares) and HI (gray bullets) listeners. Although the HI listeners showed, on average, larger bandwidths than the NH listeners by a factor of 1.2 (0.9, 1.8) [15th and 85th percentiles], the difference in bandwidths between the two groups was not significant [ $p = 0.19$ ]. This was due to the large spread of results within the group of HI listeners and the fact that

#### 4.4 Relation between cochlear response time and frequency selectivity 81

six of the HI listeners showed bandwidths within the range of the NH listeners. For one of the HI listeners, the uncertainty of the estimated ERB was considerably larger than for the other listeners. This was due to the small range of masked thresholds (8 dB) across the different notch conditions for this listener, which rendered the filter estimate less precise (cf. Tyler et al., 1984).

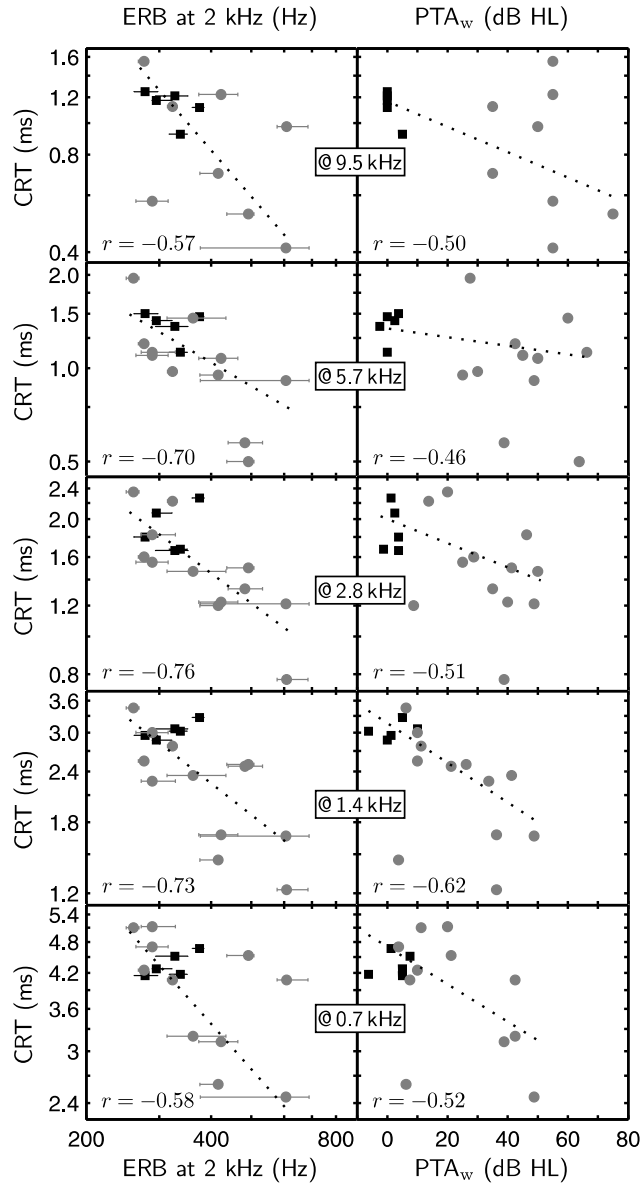
Within the group of HI listeners, the ERB at 2 kHz was significantly correlated with the individual hearing threshold at this frequency [ $r = 0.65$ ,  $p = 0.02$ ]. Here, the hearing threshold was estimated by means of a 3I-3AFC method with a 1-dB stepsize. A similar correlation was observed when the PTA (at 0.5, 1, 2, and 4 kHz) was considered instead [ $r = 0.59$ ]. The finding of a correlation between frequency selectivity and hearing threshold is consistent with previous reports in literature (e.g., Tyler et al., 1983; Moore, 1996). Typically, the correlations are less distinct for hearing losses below 30-40 dB HL (as in the present study) than for more severe losses (see Baker and Rosen, 2002).

The mean ERB for the NH listeners was 322 (SD 38) Hz. This value is larger than the value of 241 Hz predicted by the ERB function given in Glasberg and Moore (1990). Baker and Rosen (2002) found good agreement between their NH mean ERB (for 40-dB SPL 2-kHz tones) and the prediction by Glasberg and Moore. The discrepancy, observed here, may be due to variability within the NH population and the particular subset of NH listeners chosen in the present study. In view of the comparison between the HI and the NH listeners, broader filter bandwidths for the NH listeners would have resulted in more conservative estimates of significance. This may explain, why no significant bandwidth difference was found between the NH and HI listeners.

#### 4.4 Relation between CRT and frequency selectivity

Figure 4.8 (left column) shows the objective ABR-based estimates of CRTs, for the five derived bands (at 93 dB), as a function of the behaviorally derived ERBs at 2 kHz. The black squares indicate the results for the NH listeners while the results for the HI listeners are shown with the gray bullets. CRTs were estimated by subtracting the delay  $\tau_{IV+0.8}$  from the measured wave-V latencies, as discussed in section 4.2.2. The dotted regression lines were obtained by means of least trimmed squares robust

Figure 4.8: CRTs for the derived-band center frequencies of 0.7, 1.4, 2.8, 5.7, and 9.5 kHz, as a function of the ERB at 2 kHz (left column) and the  $PTA_w$  (right column). CRTs were estimated by subtracting individual  $\tau_{IV+0.8}$ -delays from the 93-dB derived-band wave-V latencies. Black squares indicate the results for the NH listeners while the results for the HI listeners are represented by gray bullets. The regression lines were obtained by means of robust regression. The errorbars to the ERBs represent the 15th and 85th percentiles, estimated via bootstrapping; see text for details.



#### 4.4 Relation between cochlear response time and frequency selectivity

83

regression (Rousseeuw, 1984). It can be seen that CRT decreased with increasing filter bandwidth, consistent with the uncertainty principle. This was significant at all frequencies [ $p < 0.05$ ], but the correlations were strongest at 1.4, 2.8 and 5.7 kHz [ $p < 0.01$ ]. The observed correlations remained largely unchanged when the results of the NH listeners were excluded and only the results of the HI listeners were considered. Also, in Fig. 4.8 (left column) it can be seen, that the inclusion of the HI listeners was crucial in order to study the relation between CRT and auditory filter bandwidth. While the HI listeners provided a relatively large span of bandwidths, the variability among the NH listeners alone would have been too small. At the frequencies of 1.4, 2.8 and 5.7 kHz, not only CRTs but also the wave-V latencies (not shown) were significantly correlated with the ERB [ $r \sim -0.57, p < 0.05$ ]. However, the correlations were stronger for the corresponding CRT estimates. The latency difference between wave V's of the 1.4- and 2.8-kHz derived bands was not significantly correlated with the ERB at 2 kHz [ $p \sim 0.09$ ].

Since the ERB was correlated with the individual hearing threshold, the correlations between CRT and ERB could have reflected an effect of hearing threshold on CRT rather than an effect of filter bandwidth per se. Therefore, the correlations between CRTs and hearing thresholds, as shown in Fig. 4.8 (right column), were examined. As suggested by Don et al. (1998), the locally weighted pure-tone average ( $PTA_w$ ) was taken as a predictor for the effect of hearing loss on the derived-band ABR. In computing the  $PTA_w$ , the audiometric threshold for the pure tone closest to the derived-band center frequency was given twice the weight of the thresholds for the two adjacent pure tones (at the highest derived-band frequency of 9.5 kHz, the audiometric threshold at 8 kHz was taken instead). As can be seen in Fig. 4.8, for all frequencies, the correlations between CRTs and the  $PTA_w$  (right column) were smaller than the corresponding correlations between CRTs and ERBs (left column). This suggests that the across-listener trend of decreasing CRT with increasing ERB indeed reflected an effect of filter bandwidth per se.

In order to test if the bandwidth estimates at 2 kHz could account for the across-listener variability in the latency data, including the latencies obtained at the lower click levels, the MEM (from section 4.2.2) was extended. In addition to the significant effects of derived-band frequency and click level, the filter bandwidth in terms of the

ERB was included, following a power-law dependence with exponent  $e$  as new model parameter:

$$\tau(f, i) = \tau_{IV+0.8} + b c^{0.93-i} f^{-d} \text{ERB}^{-e}, \quad (4.5)$$

The ERB was found to be highly significant [ $F(1, 149) = 40.6$ ,  $p < 0.0001$ ], with an estimated value of 0.6 for the exponent  $e$ . This confirmed the trend that was observed for the 93-dB latencies (Fig. 4.8, left column): Listeners with broader auditory filters at 2 kHz showed shorter derived-band latencies. Due to the inclusion of the ERB, the effect of listener group on the parameter  $b$  was no longer significant. However, the level slope  $c$  was still significantly steeper for the HI than for the NH listeners [ $p < 0.001$ ].

Also here, it was tested whether the predictive power of the ERB (for estimating CRT) was due to its correlation with the hearing threshold rather than an effect of filter bandwidth per se. The significance of the absolute hearing thresholds (in terms of the PTA and the  $\text{PTA}_w$ ) as well as the individual click thresholds was examined in a type-III ANOVA. This was done by allowing the parameter  $b$  in the MEM to be an exponential function of one of the above factors, analogous to the latency-level dependence in Eq. (4.5). For example, in case of the PTA this yielded:  $b = \exp(b_1 + b_2 \cdot \text{PTA})$ . However, none of the threshold measures reached significance [ $p > 0.15$ ], while the effect of bandwidth remained significant. Furthermore, no significant effect of age was found [ $p > 0.1$ ].

These results did not depend on the choice of  $\tau_{IV+0.8}$  as estimate of the asymptotic delay in the latency model, Eq. (4.5). The same pattern of results was obtained using a model with free parameter  $a$ . Furthermore, in the above nonlinear MEM all measured latencies (at all derived-band frequencies and click levels) were included. However, the results did not depend on a particular subset of the data: They were confirmed separately for the 93-dB data (fixed level) as well as the 1.4-kHz and 2.8-kHz derived-band data (fixed frequency).

## 4.5 Predicting frequency selectivity from derived-band ABR latencies

### 4.5.1 Background

In the preceding section, auditory filter shapes and the corresponding ERBs were measured at a single frequency and tone level. Therefore, the bandwidth results could not account for the frequency and level dependence of the ABR latencies. Nevertheless, strong correlations were observed between ERBs at 2 kHz and CRTs at the adjacent derived-band frequencies of 1.4 and 2.8 kHz (based on the 93-dB ABRs, Fig. 4.8). If such correlations between ERB and CRT were found at other frequencies, it might be possible to predict individual frequency selectivity from derived-band ABR measurements. Similar attempts have been made based on measurements of OAEs (e.g., Moleti and Sisto, 2003). The prediction of ERBs from ABR latencies could serve as an alternative to time-consuming objective tuning-curve measures of frequency selectivity in human listeners, based on masking functions of compound action potentials or ABRs (e.g., Klein and Mills, 1981; Harrison, 1984; Markessis et al., 2009). Therefore, in order to investigate, whether correlations between ERB and CRT also hold at other frequencies, additional behavioral bandwidth estimates were obtained at 1 and 4 kHz, for a subset of the listeners.

### 4.5.2 Method

In addition to the auditory filter estimates at 2 kHz, for three of the NH listeners and six of the HI listeners additional notched-noise filter estimates were obtained at 1 and 4 kHz. For the NH listeners, the 1-kHz and 4-kHz tones were presented at a fixed level of 40 dB SPL, the same level as used with the 2-kHz tones. For some of the HI listeners the levels were raised to ensure minimum sensation levels of 15 dB. The average tone levels for the HI listeners were 48 (SD 12) dB and 55 (SD 10) dB at 1 and 4 kHz, respectively. Otherwise the same stimuli and procedures as described in section 4.3.1 were used.

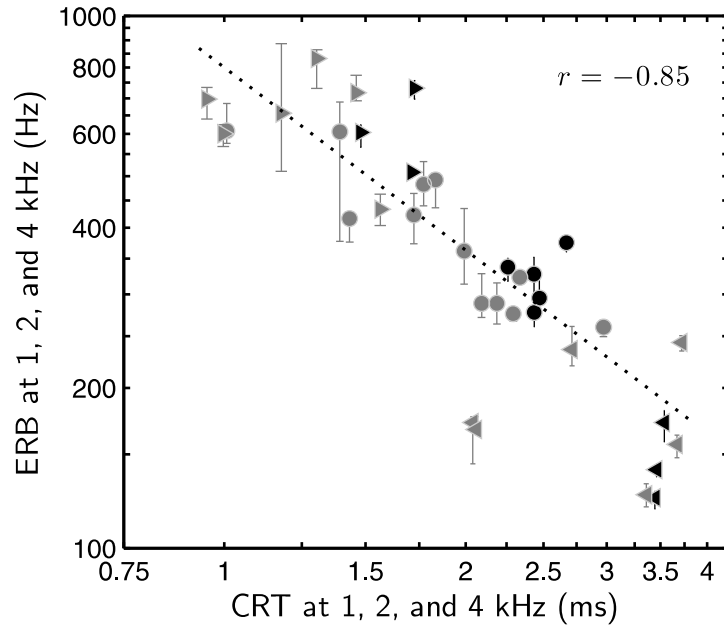


Figure 4.9: ERB at 1 kHz (left-pointing triangles), 2 kHz (bullets) and 4 kHz (right-pointing triangles) as a function of the CRT at the corresponding frequency (in double-logarithmic scaling), for the NH (black symbols) and HI listeners (gray symbols). CRTs were estimated by subtracting the individual  $\tau_{IV+0.8}$  delays from the 93-dB derived-band wave-V latencies. The regression line was obtained by means of robust regression.

### 4.5.3 Results and discussion

In order to estimate CRTs at 1, 2, and 4 kHz, first, wave-V latencies at these frequencies were predicted from the individual fits to the measured 93-dB derived-band wave-V latencies (Fig. 4.3). The CRT estimates were then obtained by subtracting the individual  $\tau_{IV+0.8}$  delays from the predicted wave-V latencies. Figure 4.9 shows the ERB at 1 kHz (left-pointing triangles), 2 kHz (bullets) and 4 kHz (right-pointing triangles) as a function of the CRT at the corresponding frequency, for the NH (black symbols) and HI listeners (gray symbols). ERBs were significantly correlated with the objective CRT estimates [ $r = -0.85$ ,  $p < 0.0001$ ], across frequency and listeners. As expected, ERBs decreased while CRTs increased with decreasing frequency.



#### 4.5 Predicting frequency selectivity from derived-band ABR latencies

87

However, while the across-listener trend of increasing CRT with decreasing ERB was apparent at 2 kHz [bullets;  $r = -0.85$ ,  $p < 0.0001$ ], no such trend was observed at 1 or 4 kHz [ $r \sim -0.2$ ; left-pointing and right-pointing triangles, respectively]. Hence, large parts of the across-listener variability in the CRTs at 1 and 4 kHz could not be accounted for in terms of the ERB at these frequencies. In other words: the individual ERBs at 1 and 4 kHz could not be reliably predicted from the ABR-based CRTs. Possible reasons for this are discussed in the following.

ERBs at 1 and 4 kHz were obtained only for nine listeners and hence, only rather strong correlations would be expected to be significant. This may partly account for the absence of correlations at 1 and 4 kHz. Furthermore, part of the unexplained variability in Fig. 4.9 might be attributable to measurement errors of the ERBs and ABR latencies. Particularly at 4 kHz, relatively small latency uncertainties can have substantial effects on the predicted ERBs (note the double-logarithmic scaling).

Changes in transport time might have obscured effects of frequency tuning on CRT at 1 and 4 kHz.<sup>3</sup> However, it remains unclear why a correlation was observed at 2 kHz, but not at 1 and 4 kHz. Ruggero and Temchin (2007) suggested that for frequencies higher than 2 kHz derived-band latencies for intense stimuli might reflect transport times rather than CRTs. This might partly account for the absent correlation at 4 kHz.

Also, it is a matter of debate whether behavioral bandwidth estimates obtained in simultaneous masking, as used in the present study, are “representative” of cochlear frequency tuning, or whether a nonsimultaneous (e.g., forward-masking) paradigm should be preferred (e.g., Moore and O’Loughlin, 1986; Shera et al., 2002; Oxenham and Shera, 2003; Ruggero and Temchin, 2005). However, the simultaneous-masking paradigm might bear a closer relationship to CRTs estimated from derived-band ABRs, which are obtained in simultaneous high-pass masking. The observed relation between CRT and ERB at 2 kHz is consistent with this hypothesis.

<sup>3</sup> Transport time, within a linear approximation, does not reflect the filter shape but rather reflects the high-frequency asymptotic slope of the filter phase response (e.g., Papoulis, 1962; Ruggero, 1980). Consequently, across-listener variability in the asymptotic phase response might have obscured effects of frequency tuning on CRT.

## 4.6 Summary

The following main observations were made in the present study. Generally, shorter derived-band latencies were observed for the HI than for the NH listeners, although prolonged latencies were obtained for some of the HI listeners at low click presentation levels. This seemed to be attributable to the low click sensation levels for these listeners. Behaviorally derived auditory-filter bandwidths at 2 kHz accounted for part of the across-listener variability in the ABR latencies, consistent with the expectation that CRT decreases with increasing filter bandwidth. Additional behavioral filter measurements at 1 and 4 kHz in a subset of the listeners indicated, that correlations between CRTs and bandwidths at these frequencies may not be as clear as the ones observed at 2 kHz and with a larger number of listeners. Hence, it remains to be seen whether individual bandwidths can be predicted from derived-band ABR latencies with reasonable precision.

# 5

---

## General discussion

---

In this thesis, behavioral and objective noninvasive methods were used to study peripheral auditory processing in listeners with normal and impaired hearing. One objective was to identify, or at least narrow down, actual sites and mechanisms of hearing impairment in the auditory periphery. Another objective was to explore possible consequences of such impairments for speech reception in background noise.

The behavioral experiments presented in chapter 2 revealed deficits in frequency selectivity and phase-locking-based TFS processing in HI listeners, in low-frequency regions of normal hearing. The observed deficits in frequency selectivity and TFS processing were not correlated. Hence, they seemed to be due to independent impairment factors, such as, for example, OHC and IHC loss, respectively. Monaural and binaural TFS deficits, however, were found to be related, suggesting that the binaural deficits might have been caused by a monaural impairment factor. Regarding speech reception, SRTs in lateralized noise and a two-talker background, but not in amplitude-modulated noise, were correlated with TFS-processing performance. This suggests that the auditory system might utilize TFS cues in order to accomplish spatial segregation and talker separation, two important aspects of the cocktail party problem. Hence, indeed, TFS deficits might, to a certain degree, account for the problems of HI people with understanding speech in complex listening situations. The two listeners with an obscure dysfunction showed deficits in frequency selectivity and in binaural, but not monaural, TFS processing. The latter finding may suggest an impairment mechanism at or later than the stage of binaural integration in these listeners. This impairment mechanism might also be responsible for their poor reception of lowpass-filtered speech, which, in turn, may account for part of their perceived difficulty with understanding speech in noisy backgrounds.

As discussed in section 2.9, one possible explanation for deficits in TFS pro-

cessing might be distortions of the spatiotemporal cochlear response in HI listeners. Therefore, two further chapters dealt specifically with CRT. Combining behavioral and objective paradigms, chapter 3 demonstrated a link between early cochlear disparities and perceived lateralization of tones that were interaurally mismatched in frequency. The behavioral estimates of CRT and TW velocity were consistent with estimates derived from ABR measurements for the same NH listeners. However, the behavioral paradigm proved to be applicable only at low to mid frequencies. Therefore, chapter 4 examined CRTs for NH and HI listeners using ABR measurements, to investigate possible alterations of CRTs due to hearing impairment within a broader range of frequencies. Shorter latencies were observed for the HI listeners than for the NH listeners. In order to test whether this could be attributed to reduced frequency selectivity in the HI listeners, individual auditory-filter bandwidths were estimated behaviorally. A relationship between CRT and frequency selectivity was observed: across listeners, CRT decreased with increasing filter bandwidth. This illustrated that a larger-than-normal across-listener variability within the population of HI listeners can provide a means for studying relationships between auditory functions, for which the span of results among NH listeners alone would be too small. The observed relationship suggests that changes in cochlear frequency tuning in HI listeners may result in alterations of CRTs and thereby in changes of the spatiotemporal cochlear response. However, changes in the spatiotemporal response pattern might not be reflected fully by changes in frequency selectivity at a single frequency. Hence, it cannot be ruled out that the TFS deficits observed for the HI listeners in chapter 2 were due to changes in the spatiotemporal response pattern.

In order to gain further insight into the role of spatiotemporal processing in NH and HI human listeners, future studies could investigate relations between individual estimates of CRT (disparities) and performance in other behavioral tasks that have been discussed in the context of spatiotemporal processing, such as pitch perception and lateralization. The behavioral paradigm presented in chapter 3 might provide valuable information about CRT disparities, particularly for frequencies below 500 Hz, where the accuracy of objective methods is limited. Also, modeling of basic auditory functions might provide valuable insights into peripheral impairment mechanisms. The findings presented in chapter 2 on auditory deficits, as well as preserved

auditory abilities, could serve as constraints for such models of the impaired auditory system. An example is the observation that background noise did not have a larger effect on TFS processing for the HI than for the NH listeners: although the acuity of TFS processing was decreased for the HI listeners, it seemed to be as robust to noise interference as for the NH listeners. In addition to such a top-down approach, it would be interesting to model the effect of distortions of the cochlear response in the framework of spatiotemporal models (cf. Carney, 1994). Interestingly, several algorithms employed in computational auditory scene analysis also rely on TFS information, in terms of correlogram or cross-correlogram analyses based on coincidence detection (Wang and Brown, 2006).

The TFS deficits observed in the HI listeners (chapter 2) were neither correlated with absolute hearing thresholds nor with frequency selectivity. However, they bore a relationship to speech reception performance in noise. Hence, it might be important to take measures of TFS-processing abilities into account when defining an “auditory profile” for listeners with impaired hearing. Such profiles aim to efficiently characterize an individual’s communication handicap, in order to maximize their benefit from assistive devices. Although it might be impossible to tackle TFS deficits directly, knowledge about the presence of such deficits might have implications for compensation strategies, such as hearing aids (Moore, 2008b). For example, an individual with little ability to process TFS information will rely mainly on envelope cues, and consequently these should be preserved as much as possible. On the other hand, hearing aids should provide the TFS with high fidelity to individuals who retain some ability to process TFS. Also, it has been suggested that minimal TFS information is conveyed by current cochlear implant systems (see Wilson and Dorman, 2008). The present findings on a potential contribution of TFS cues to speech understanding in complex listening situations further encourage attempts to convey TFS information to cochlear implantees, for example by means of combined electric and acoustic stimulation paradigms. Such advances should help to alleviate problems that people with hearing difficulties face in their everyday life.



---

## Bibliography

---

- Abbas, P. J. (1981). Auditory-nerve fiber responses to tones in a noise masker. *Hear. Res.*, **5**(1), 69–80.
- Baker, R. J. and Rosen, S. (2002). Auditory filter nonlinearity in mild/moderate hearing impairment. *J. Acoust. Soc. Am.*, **111**(3), 1330–1339.
- Buss, E., Hall, J. W., and Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear Hear.*, **25**(3), 242–250.
- Buus, S., Scharf, B., and Florentine, M. (1984). Lateralization and frequency selectivity in normal and impaired hearing. *J. Acoust. Soc. Am.*, **76**(1), 77–86.
- Carney, L. H. (1994). Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness. *Hear. Res.*, **76**(1-2), 31–44.
- Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., and Colburn, H. S. (2002). Auditory phase opponency: A temporal model for masked detection at low frequencies. *Acta. Acust. Acust.*, **88**(3), 334–347.
- Chen, H. and Zeng, F.-G. (2004). Frequency modulation detection in cochlear implant subjects. *J. Acoust. Soc. Am.*, **116**(4), 2269–2277.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.*, **25**(5), 975–979.
- Clarey, J. C., Barone, P., and Imig, T. J. (1992). Physiology of thalamus and cortex. In A. N. Popper and R. R. Fay (Eds.), *The Mammalian Auditory Pathway: Neurophysiology* (pp. 232–334). Springer-Verlag, New York.
- Colburn, H. S. (1996). Computational models of binaural processing. In H. Hawkins, T. McMullen, A. Popper, and R. Fay (Eds.), *Auditory Computation*. Springer-Verlag, New York.

- Colburn, H. S. and Häusler, R. (1980). Note on the modeling of binaural interaction in impaired auditory systems. In G. van den Brink and F. A. Bilsen (Eds.), *Psychophysical, Physiological, and Behavioral Studies in Hearing*. Delft University Press, Delft.
- Costalupes, J. A. (1985). Representation of tones in noise in the responses of auditory nerve fibers in cats. I. Comparison with detection thresholds. *J. Neurosci.*, **5**(12), 3261–3269.
- Davis, A. C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *Int. J. Epidemiol.*, **18**(4), 911–917.
- Deatherage, B. H. and Hirsh, I. J. (1959). Auditory localization of clicks. *J. Acoust. Soc. Am.*, **31**(4), 486–492.
- Demany, L. and Semal, C. (1989). Detection thresholds for sinusoidal frequency modulation. *J. Acoust. Soc. Am.*, **85**(3), 1295–1301.
- Deng, L. and Geisler, C. D. (1987). A composite auditory model for processing speech sounds. *J. Acoust. Soc. Am.*, **82**(6), 2001–2012.
- Don, M. and Eggermont, J. J. (1978). Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *J. Acoust. Soc. Am.*, **63**(4), 1084–1092.
- Don, M. and Elberling, C. (1994). Evaluating residual background noise in human auditory brain-stem responses. *J. Acoust. Soc. Am.*, **96**(5), 2746–2757.
- Don, M. and Kwong, B. (2002). Auditory brainstem response: Differential diagnosis. In J. Katz (Ed.), *Handbook of Clinical Audiology* (pp. 274–297). Lippincott Williams & Wilkins, Philadelphia.
- Don, M., Ponton, C. W., Eggermont, J. J., and Kwong, B. (1998). The effects of sensory hearing loss on cochlear filter times estimated from auditory brainstem response latencies. *J. Acoust. Soc. Am.*, **104**(4), 2280–2289.
- Don, M., Ponton, C. W., Eggermont, J. J., and Masuda, A. (1993). Gender differences in cochlear response time: An explanation for gender amplitude differences in the unmasked auditory brain-stem response. *J. Acoust. Soc. Am.*, **94**(4), 2135–2148.
- Donaldson, G. S. and Ruth, R. A. (1993). Derived band auditory brain-stem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects. *J. Acoust. Soc. Am.*, **93**(2), 940–951.



- Donaldson, G. S. and Ruth, R. A. (1996). Derived-band auditory brain-stem response estimates of traveling wave velocity in humans: II. Subjects with noise-induced hearing loss and Meniere's disease. *J. Speech Hear. Res.*, **39**(3), 534–545.
- Dreschler, W. A. and Plomp, R. (1985). Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *J. Acoust. Soc. Am.*, **78**(4), 1261–1270.
- Durlach, N. I., Thompson, C. L., and Colburn, H. S. (1981). Binaural interaction in impaired listeners. A review of past research. *Audiology*, **20**(3), 181–211.
- Eggermont, J. J. (1976). Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea. *J. Acoust. Soc. Am.*, **60**(5), 1132–1139.
- Eggermont, J. J. (1979). Narrow-band AP latencies in normal and recruiting human ears. *J. Acoust. Soc. Am.*, **65**(2), 463–470.
- Eggermont, J. J. and Don, M. (1980). Analysis of the click-evoked brainstem potentials in humans using high-pass noise masking. II. Effect of click intensity. *J. Acoust. Soc. Am.*, **68**(6), 1671–1675.
- Elberling, C. and Don, M. (1984). Quality estimation of averaged auditory brainstem responses. *Scand. Audiol.*, **13**(3), 187–197.
- Elberling, C. and Wahlgreen, O. (1985). Estimation of auditory brainstem response, ABR, by means of Bayesian inference. *Scand. Audiol.*, **14**(2), 89–96.
- Evans, E. F. and Harrison, R. V. (1976). Correlation between cochlear outer hair cell damage and deterioration of cochlear nerve tuning properties in the guinea-pig. *J. Physiol.*, **256**(1), 43P–44P.
- Festen, J. M. and Plomp, R. (1983). Relations between auditory functions in impaired hearing. *J. Acoust. Soc. Am.*, **73**(2), 652–662.
- Freyman, R. L. and Nelson, D. A. (1991). Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners. *J. Speech Hear. Res.*, **34**(6), 1371–1386.
- Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.*, **211**(1-2), 74–84.
- Gabriel, K. J., Koehnke, J., and Colburn, H. S. (1992). Frequency dependence of binaural performance in listeners with impaired binaural hearing. *J. Acoust. Soc. Am.*, **91**(1), 336–347.

- Geisler, C. D. and Sinex, D. G. (1982). Responses of primary auditory fibers to brief tone bursts. *J. Acoust. Soc. Am.*, **72**(3), 781–794.
- Geisler, C. D. and Sinex, D. G. (1983). Comparison of click responses of primary auditory fibers with minimum-phase predictions. *J. Acoust. Soc. Am.*, **73**(5), 1671–1675.
- Gilbert, G. and Lorenzi, C. (2006). The ability of listeners to use recovered envelope cues from speech fine structure. *J. Acoust. Soc. Am.*, **119**(4), 2438–2444.
- Glasberg, B. R. and Moore, B. C. J. (1989). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scand. Audiol., Suppl.*, **32**, 1–25.
- Glasberg, B. R. and Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hear. Res.*, **47**(1-2), 103–138.
- Gnansia, D., Jourdes, V., and Lorenzi, C. (2008). Effect of masker modulation depth on speech masking release. *Hear. Res.*, **239**(1-2), 60–68.
- Goldstein, J., Baer, T., and Kiang, N. (1971). A theoretical treatment of latency, group delay and tuning characteristics for auditory-nerve responses to clicks and tones. In M. Sachs (Ed.), *The Physiology of the Auditory System* (pp. 133–141). National Educational Consultants, Baltimore.
- Gorga, M. P., Kaminski, J. R., Beauchaine, K. A., and Jesteadt, W. (1988). Auditory brainstem responses to tone bursts in normally hearing subjects. *J. Speech. Hear. Res.*, **31**(1), 87–97.
- Grant, K. W. (1987). Frequency modulation detection by normally hearing and profoundly hearing-impaired listeners. *J. Speech Hear. Res.*, **30**(4), 558–563.
- Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *J. Acoust. Soc. Am.*, **33**, 1344–1356.
- Hall, J. W., Tyler, R. S., and Fernandes, M. A. (1984). Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners. *J. Speech Hear. Res.*, **27**(1), 145–154.
- Harrison, R. V. (1984). Objective measures of cochlear frequency selectivity in animals and in man. A review. *Acta Neurol. Belg.*, **84**(5), 213–232.
- Häusler, R., Colburn, S., and Marr, E. (1983). Sound localization in subjects with impaired hearing. Spatial-discrimination and interaural-discrimination tests. *Acta Oto-Laryngol., Suppl.*, **400**, 1–62.

- Hawkins, D. B. and Wightman, F. L. (1980). Interaural time discrimination ability of listeners with sensorineural hearing loss. *Audiology*, **19**(6), 495–507.
- Heil, P. (2004). First-spike latency of auditory neurons revisited. *Curr. Opin. Neurobiol.*, **14**(4), 461–467.
- Heil, P. and Neubauer, H. (2003). A unifying basis of auditory thresholds based on temporal summation. *Proc. Natl. Acad. Sci. U.S.A.*, **100**(10), 6151–6156.
- Hinchcliffe, R. (1992). King-Kopetzky syndrome: An auditory stress disorder? *J. Audiol. Med.*, **1**, 89–98.
- Hopkins, K. and Moore, B. C. J. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J. Acoust. Soc. Am.*, **122**(2), 1055–1068.
- Hopkins, K. and Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J. Acoust. Soc. Am.*, **125**(1), 442–446.
- Hopkins, K., Moore, B. C. J., and Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J. Acoust. Soc. Am.*, **123**(2), 1140–1153.
- Horst, J. W. (1987). Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects. *J. Acoust. Soc. Am.*, **82**(3), 874–885.
- Horwitz, A. R., Dubno, J. R., and Ahlstrom, J. B. (2002). Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. *J. Acoust. Soc. Am.*, **111**(1), 409–416.
- ISO 389-8 (2004). Acoustics—Reference zero for the calibration of audiometric equipment—Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones. International Organization for Standardization, Geneva.
- Jesteadt, W. (1980). An adaptive procedure for subjective judgments. *Percept. Psychophys.*, **28**(1), 85–88.
- Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J. Neurophysiol.*, **71**(3), 1022–1036.

- Joris, P. X., de Sande, B. V., Louage, D. H., and van der Heijden, M. (2006). Binaural and cochlear disparities. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(34), 12917–12922.
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Percept. Psychophys.*, **49**(3), 227–229.
- King, K. and Stephens, D. (1992). Auditory and psychological factors in ‘auditory disability with normal hearing’. *Scand. Audiol.*, **21**(2), 109–114.
- Kinkel, M., Holube, I., and Kollmeier, B. (1988). Zusammenhang verschiedener Parameter binauralen Hörens bei Schwerhörigen (Relation between parameters of binaural hearing in hearing impaired subjects). In *Fortschritte der Akustik - DAGA 1988*, (pp. 629–632). DPG Kongreß-GmbH, Bad Honnef.
- Klein, A. J. and Mills, J. H. (1981). Physiological (waves I and V) and psychophysical tuning curves in human subjects. *J. Acoust. Soc. Am.*, **69**(3), 760–768.
- Koehnke, J., Culotta, C. P., Hawley, M. L., and Colburn, H. S. (1995). Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing. *Ear Hear.*, **16**(4), 331–353.
- Kunov, H. and Abel, S. M. (1981). Effects of rise/decay time on the lateralization of interaurally delayed 1-kHz tones. *J. Acoust. Soc. Am.*, **69**(3), 769–773.
- Lacher-Fougère, S. and Demany, L. (1998). Modulation detection by normal and hearing-impaired listeners. *Audiology*, **37**(2), 109–121.
- Lacher-Fougère, S. and Demany, L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *J. Acoust. Soc. Am.*, **118**(4), 2519–2526.
- Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**(4), 963–974.
- Levitt, H. and Rabiner, L. R. (1967). Binaural release from masking for speech and gain in intelligibility. *J. Acoust. Soc. Am.*, **42**(3), 601–608.
- Lieberman, M. C. and Dodds, L. W. (1984). Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear. Res.*, **16**(1), 55–74.
- Lindstrom, M. L. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, **46**(3), 673–687.
- Loeb, G. E., White, M. W., and Merzenich, M. M. (1983). Spatial cross-correlation. A proposed mechanism for acoustic pitch perception. *Biol. Cybern.*, **47**(3), 149–163.

- Lorenzi, C., Debrulle, L., Garnier, S., Fleuriot, P., and Moore, B. C. J. (2009). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *J. Acoust. Soc. Am.*, **125**(1), 27–30.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(49), 18866–18869.
- Lorenzi, C. and Moore, B. C. J. (2008). Role of temporal envelope and fine structure cues in speech perception: A review. In T. Dau, J. M. Buchholz, J. M. Harte, and T. U. Christiansen (Eds.), *Auditory Signal Processing in Hearing-Impaired Listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007)*. Centertryk, Denmark.
- Magezi, D. A. and Krumbholz, K. (2008). Can the binaural system extract fine-structure interaural time differences from noncorresponding frequency channels? *J. Acoust. Soc. Am.*, **124**(5), 3095.
- Markessis, E., Poncelet, L., Colin, C., Coppens, A., Hoonhorst, I., Kadhim, H., and Deltenre, P. (2009). Frequency tuning curves derived from auditory steady state evoked potentials: A proof-of-concept study. *Ear Hear.*, **30**(1), 43–53.
- Middelweerd, M. J., Festen, J. M., and Plomp, R. (1990). Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram. *Audiology*, **29**(1), 1–7.
- Miller, R. L., Schilling, J. R., Franck, K. R., and Young, E. D. (1997). Effects of acoustic trauma on the representation of the vowel "eh" in cat auditory nerve fibers. *J. Acoust. Soc. Am.*, **101**(6), 3602–3616.
- Moleti, A. and Sisto, R. (2003). Objective estimates of cochlear tuning by otoacoustic emission analysis. *J. Acoust. Soc. Am.*, **113**(1), 423–429.
- Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear Hear.*, **17**(2), 133–161.
- Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing*, (pp. 197–204). Academic Press, San Diego, CA.
- Moore, B. C. J. (2008a). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Oto-Laryngol.*, **9**(4), 399–406.

- Moore, B. C. J. (2008b). The choice of compression speed in hearing aids: Theoretical and practical considerations, and the role of individual differences. *Trends in Amplification*, **12**(2), 102–112.
- Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *J. Audio Engin. Soc.*, **45**(4), 224–240.
- Moore, B. C. J., Glasberg, B. R., and Hopkins, K. (2006). Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hear. Res.*, **222**(1-2), 16–27.
- Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1985). Relative dominance of individual partials in determining the pitch of complex tones. *J. Acoust. Soc. Am.*, **77**(5), 1853–1860.
- Moore, B. C. J. and O’Loughlin, B. J. (1986). The use of nonsimultaneous masking to measure frequency selectivity and suppression. In *Frequency Selectivity in Hearing*. Academic Press, London.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1990). Auditory filter shapes at low center frequencies. *J. Acoust. Soc. Am.*, **88**(1), 132–140.
- Moore, B. C. J. and Sek, A. (1996). Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *J. Acoust. Soc. Am.*, **100**(4), 2320–2331.
- Moore, B. C. J. and Skrodzka, E. (2002). Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation. *J. Acoust. Soc. Am.*, **111**, 327–335.
- Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *J. Acoust. Soc. Am.*, **106**(5), 2761–2778.
- Narula, A. A. and Mason, S. M. (1988). Selective dysacusis—a preliminary report. *J. R. Soc. Med.*, **81**(6), 338–340.
- Neely, S. T., Norton, S. J., Gorga, M. P., and Jesteadt, W. (1988). Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli. *J. Acoust. Soc. Am.*, **83**(2), 652–656.
- Nie, K., Stickney, G., and Zeng, F.-G. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Trans. Biomed. Eng.*, **52**(1), 64–73.

- Noordhoek, I. M., Houtgast, T., and Festen, J. M. (2001). Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. *J. Acoust. Soc. Am.*, **109**(3), 1197–1212.
- Norton, S. J. and Neely, S. T. (1987). Tone-burst-evoked otoacoustic emissions from normal-hearing subjects. *J. Acoust. Soc. Am.*, **81**(6), 1860–1872.
- Oxenham, A. J. and Shera, C. A. (2003). Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J. Assoc. Res. Oto-Laryngol.*, **4**(4), 541–554.
- Oxenham, A. J. and Simonson, A. M. (2009). Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J. Acoust. Soc. Am.*, **125**(1), 457–468.
- Papoulis, A. (1962). *The Fourier Integral and Its Applications*. McGraw-Hill, New York.
- Parker, D. J. and Thornton, A. R. (1978a). The validity of the derived cochlear nerve and brainstem evoked responses of the human auditory system. *Scand. Audiol.*, **7**(1), 45–52.
- Parker, D. J. and Thornton, A. R. (1978b). Frequency specific components of the cochlear nerve and brainstem evoked responses of the human auditory system. *Scand. Audiol.*, **7**(1), 53–60.
- Patterson, R. D. and Moore, B. C. J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B. C. J. Moore (Ed.), *Frequency Selectivity in Hearing*. Academic Press, London.
- Patterson, R. D. and Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *J. Acoust. Soc. Am.*, **67**(1), 229–245.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *J. Acoust. Soc. Am.*, **72**(6), 1788–1803.
- Peters, R. W., Moore, B. C., and Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.*, **103**(1), 577–587.
- Peterson, G. and Barney, H. (1952). Control methods used in a study of the vowels. *J. Acoust. Soc. Am.*, **24**, 175–184.

- Pichora-Fuller, M. K. and Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *J. Acoust. Soc. Am.*, **91**(4), 2129–2135.
- Pinheiro, J. and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag, New York.
- Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.*, **63**(2), 533–549.
- Ponton, C. W., Eggermont, J. J., Coupland, S. G., and Winkelaar, R. (1992). Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs). *J. Acoust. Soc. Am.*, **91**(3), 1576–1586.
- Qin, M. K. and Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.*, **114**(1), 446–454.
- Recio, A. and Rhode, W. S. (2000). Basilar membrane responses to broadband stimuli. *J. Acoust. Soc. Am.*, **108**(5), 2281–2298.
- Rhode, W. S., Geisler, C. D., and Kennedy, D. T. (1978). Auditory nerve fiber response to wide-band noise and tone combinations. *J. Neurophysiol.*, **41**(3), 692–704.
- Robles, L. and Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiol. Rev.*, **81**(3), 1305–1352.
- Rosen, S. (1987). Phase and the hearing-impaired. In M. E. H. Schouten (Ed.), *The Psychophysics of Speech Perception*. Springer-Verlag, New York.
- Rosen, S., Baker, R. J., and Darling, A. (1998). Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *J. Acoust. Soc. Am.*, **103**(5), 2539–2550.
- Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (2007). Aging in binaural hearing begins in mid-life: Evidence from cortical auditory-evoked responses to changes in interaural phase. *J. Neurosci.*, **27**(42), 11172–11178.
- Rousseeuw, P. J. (1984). Least median of squares regression. *J. Am. Stat. Assoc.*, **79**, 871–880.
- Ruggero, M. A. (1980). Systematic errors in indirect estimates of basilar membrane travel times. *J. Acoust. Soc. Am.*, **67**(2), 707–710.



- Ruggero, M. A. (1992). Physiology and coding of sound in the auditory nerve. In A. N. Popper and R. R. Fay (Eds.), *The Mammalian Auditory Pathway: Neurophysiology* (pp. 34–93). Springer-Verlag, New York.
- Ruggero, M. A. (1994). Cochlear delays and traveling waves: Comments on ‘Experimental look at cochlear mechanics’. *Audiology*, **33**(3), 131–142.
- Ruggero, M. A. and Rich, N. C. (1987). Timing of spike initiation in cochlear afferents: dependence on site of innervation. *J. Neurophysiol.*, **58**(2), 379–403.
- Ruggero, M. A. and Temchin, A. N. (2005). Unexceptional sharpness of frequency tuning in the human cochlea. *Proc. Natl. Acad. Sci. U.S.A.*, **102**(51), 18614–18619.
- Ruggero, M. A. and Temchin, A. N. (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods. *J. Assoc. Res. Oto-Laryngol.*, **8**(2), 153–166.
- Rutten, W. L. (1986). The influence of cochlear hearing loss and probe tone level on compound action potential tuning curves in humans. *Hear. Res.*, **21**(3), 195–204.
- Saberi, K. (1995). Some considerations on the use of adaptive methods for estimating interaural-delay thresholds. *J. Acoust. Soc. Am.*, **98**(3), 1803–1806.
- Santurette, S. and Dau, T. (2007). Binaural pitch perception in normal-hearing and hearing-impaired listeners. *Hear. Res.*, **223**(1-2), 29–47.
- Saunders, G. H. and Haggard, M. P. (1989). The clinical assessment of obscure auditory dysfunction—1. Auditory and psychological factors. *Ear Hear.*, **10**(3), 200–208.
- Saunders, G. H. and Haggard, M. P. (1992). The clinical assessment of ‘Obscure Auditory Dysfunction’ (OAD) 2. Case control analysis of determining factors. *Ear Hear.*, **13**(4), 241–254.
- Scharf, B. (1972). Frequency selectivity and sound localization. In B. L. Cardozo (Ed.), *Symposium on Hearing Theory* (pp. 115–122). IPO, Eindhoven.
- Scharf, B., Florentine, M., and Meiselman, C. (1976). Critical band in auditory lateralization. *Sensory processes*, **1**, 109–126.
- Schneider, B. A., Daneman, M., and Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Can. J. Exp. Psychol.*, **56**(3), 139–152.
- Schroeder, M. R. (1977). New viewpoints in binaural interaction. In E. F. Evans and J. P. Wilson (Eds.), *Psychophysics and Physiology of Hearing* (pp. 455–467). Academic Press, New York.

- Schubert, E. D. and Elpern, B. S. (1959). Psychophysical estimate of the velocity of the traveling wave. *J. Acoust. Soc. Am.*, **31**(7), 990–994.
- Schubert, E. D. and Schultz, M. C. (1962). Some aspects of binaural signal selection. *J. Acoust. Soc. Am.*, **34**(6), 844–849.
- Schuknecht, H. F. and Woellner, R. C. (1953). Hearing losses following partial section of the cochlear nerve. *Laryngoscope*, **63**(6), 441–465.
- Shamma, S. (2001). On the role of space and time in auditory processing. *Trends Cogn. Sci.*, **5**(8), 340–348.
- Shamma, S. and Klein, D. (2000). The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *J. Acoust. Soc. Am.*, **107**(5), 2631–2644.
- Shamma, S. A. (1985). Speech processing in the auditory system. II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am.*, **78**(5), 1622–1632.
- Shamma, S. A., Shen, N. M., and Gopalaswamy, P. (1989). Stereausis: Binaural processing without neural delays. *J. Acoust. Soc. Am.*, **86**(3), 989–1006.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, **270**(5234), 303–304.
- Shera, C. A., Guinan, J. J., and Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Natl. Acad. Sci. U.S.A.*, **99**(5), 3318–3323.
- Smoski, W. J. and Trahiotis, C. (1986). Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. *J. Acoust. Soc. Am.*, **79**(5), 1541–1547.
- Staffel, J. G., Hall, J. W., Grose, J. H., and Pillsbury, H. C. (1990). NoSo and NoS $\pi$  detection as a function of masker bandwidth in normal-hearing and cochlear-impaired listeners. *J. Acoust. Soc. Am.*, **87**(4), 1720–1727.
- Stern, M. and Trahiotis, C. (1995). Models of binaural interaction. In B. C. J. Moore (Ed.), *Hearing*. Academic Press, San Diego, CA.
- Strelcyk, O., Christoforidis, D., and Dau, T. (2009). Relation between derived-band auditory brainstem response latencies and frequency selectivity. *J. Acoust. Soc. Am.* (submitted).

## Bibliography

105

- Strelcyk, O. and Dau, T. (2009a). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J. Acoust. Soc. Am.* (in press).
- Strelcyk, O. and Dau, T. (2009b). Estimation of cochlear response times using lateralization of frequency-mismatched tones. *J. Acoust. Soc. Am.* (submitted).
- Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (1998). Temporal processing in the aging auditory system. *J. Acoust. Soc. Am.*, **104**(4), 2385–2399.
- Terhardt, E. (1974). Pitch, consonance, and harmony. *J. Acoust. Soc. Am.*, **55**(5), 1061–1069.
- Thornton, A. R. and Farrell, G. (1991). Apparent travelling wave velocity changes in cases of endolymphatic hydrops. *Scand. Audiol.*, **20**(1), 13–18.
- Tognola, G., Grandori, F., and Ravazzani, P. (1997). Time-frequency distributions of click-evoked otoacoustic emissions. *Hear. Res.*, **106**(1-2), 112–122.
- Turner, C. W. (1987). Effects of noise and hearing loss upon frequency discrimination. *Audiology*, **26**(3), 133–140.
- Turner, C. W. and Nelson, D. A. (1982). Frequency discrimination in regions of normal and impaired sensitivity. *J. Speech Hear. Res.*, **25**(1), 34–41.
- Tyler, R. S., Hall, J. W., Glasberg, B. R., Moore, B. C., and Patterson, R. D. (1984). Auditory filter asymmetry in the hearing impaired. *J. Acoust. Soc. Am.*, **76**(5), 1363–1368.
- Tyler, R. S., Wood, E. J., and Fernandes, M. (1983). Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.*, **74**(4), 1190–1199.
- van Schijndel, N. H., Houtgast, T., and Festen, J. M. (2001). Effects of degradation of intensity, time, or frequency content on speech intelligibility for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.*, **110**(1), 529–542.
- von Békésy, G. (1933). Über den Knall und die Theorie des Hörens (Clicks and the theory of hearing). *Physik. Z.*, **34**, 577–582.
- von Békésy, G. (1963a). Hearing theories and complex sounds. *J. Acoust. Soc. Am.*, **35**(4), 588–601.
- von Békésy, G. (1963b). Three experiments concerned with pitch perception. *J. Acoust. Soc. Am.*, **35**(4), 602–606.

- Wagener, K., Josvassen, J. L., and Ardenkjaer, R. (2003). Design, optimization and evaluation of a Danish sentence test in noise. *Int. J. Audiol.*, **42**(1), 10–17.
- Wallhagen, M. I., Strawbridge, W. J., Cohen, R. D., and Kaplan, G. A. (1997). An increasing prevalence of hearing impairment and associated risk factors over three decades of the Alameda County Study. *Am. J. Public Health*, **87**(3), 440–442.
- Wang, D. L. and Brown, G. J. (2006). Fundamentals of computational auditory scene analysis. In D. L. Wang and G. J. Brown (Eds.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. John Wiley & Sons, Inc., New Jersey.
- Wang, J., Powers, N. L., Hofstetter, P., Trautwein, P., Ding, D., and Salvi, R. (1997). Effects of selective inner hair cell loss on auditory nerve fiber threshold, tuning and spontaneous and driven discharge rate. *Hear. Res.*, **107**(1-2), 67–82.
- Wightman, F. L. and Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am.*, **91**(3), 1648–1661.
- Wilson, B. S. and Dorman, M. F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hear. Res.*, **242**(1-2), 3–21.
- Woolf, N. K., Ryan, A. F., and Bone, R. C. (1981). Neural phase-locking properties in the absence of cochlear outer hair-cells. *Hear. Res.*, **4**(3-4), 335–346.
- World Health Organization (1997). Prevention of noise-induced hearing loss. Report of an informal consultation. Geneva.
- World Health Organization (2004). Guidelines for hearing aids and services for developing countries, 2nd edition. Geneva.
- Yost, W. A. (1974). Discriminations of interaural phase differences. *J. Acoust. Soc. Am.*, **55**(6), 1299–1303.
- Zeng, F.-G., Nie, K., Liu, S., Stickney, G., Rio, E. D., Kong, Y.-Y., and Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues. *J. Acoust. Soc. Am.*, **116**(3), 1351–1354.
- Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., and Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. U.S.A.*, **102**(7), 2293–2298.
- Zerlin, S. (1969). Traveling-wave velocity in the human cochlea. *J. Acoust. Soc. Am.*, **46**(4), 1011–1015.

## Bibliography

107

- Zhao, F. and Stephens, D. (2000). Subcategories of patients with King-Kopetzky syndrome. *Br. J. Audiol.*, **34**(4), 241–256.
- Zhao, F. and Stephens, D. (2006). Distortion product otoacoustic emissions in patients with King-Kopetzky syndrome. *Int. J. Audiol.*, **45**(1), 34–39.
- Zurek, P. M. (1993). A note on onset effects in binaural hearing. *J. Acoust. Soc. Am.*, **93**(2), 1200–1201.
- Zurek, P. M. and Formby, C. (1981). Frequency-discrimination ability of hearing-impaired listeners. *J. Speech Hear. Res.*, **24**(1), 108–112.
- Zwicker, E. (1956). Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The elementary foundations for determining the information capacity of the auditory system). *Acustica*, **6**, 365–381.

One of the most common complaints of people with impaired hearing concerns their difficulty with understanding speech. Particularly in the presence of background noise, hearing-impaired people often encounter great difficulties with speech communication. In most cases, the benefit from hearing aids varies among listeners. It has been hypothesized that part of the difficulty arises from changes in the perception of sounds that are well above hearing threshold, such as reduced frequency selectivity and deficits in the processing of temporal-fine-structure at the output of the inner-ear filters. Here, relations between frequency selectivity, temporal fine-structure processing, and speech reception were investigated in listeners with normal and impaired hearing, using behavioral listening experiments as well as objective measurements of auditory evoked potentials. This work provides insights into factors affecting auditory processing in listeners with impaired hearing and may have implications for future models of impaired auditory signal processing as well as advanced compensation strategies.

## **DTU Electrical Engineering**

Department of Electrical Engineering

---

Ørsted's Plads  
Building 348  
DK-2800 Kgs. Lyngby  
Denmark  
Tel: (+45) 45 25 38 00  
Fax: (+45) 45 93 16 34  
[www.elektro.dtu.dk](http://www.elektro.dtu.dk)

ISBN 978-87-92465-02-3